
STUDIES OF ANGIOSPERM PHYLOGENY USING PROTEIN SEQUENCES¹

P. G. Martin² and J. M. Dowd²

ABSTRACT

In previous papers we have reported the N-terminal 40 amino acids of the small subunit of rubisco for samples from four families of gymnosperms, nine families of monocotyledons, and 26 families of dicotyledons. We expanded this list to 122 families of dicots and derived a phylogenetic tree for all 335 species. The main computing program used was HENNIG86, with which a reliable result can be assured with only 17 taxa or less, so a major part of this paper is concerned with the strategy adopted to divide the 335 species and then to build the parts into an overall tree that is as accurate and objective as possible. Comparison with other taxonomy suggests that, at the level of placing genera into families, our methods give results that are at least 90% accurate. At higher taxonomic levels accuracy may decrease, and the result should be regarded not as a firm conclusion but as a working hypothesis for subsequent testing using the longer sequences from nucleic acids. Topics discussed include heterogeneity within species, the nature of the N-terminus of rubisco-SSU, and evidence that natural selection is powerful in determining amino acid sequence. The rate of evolution has been shown to vary between major taxa, and data suggest that angiosperms originated in the Jurassic.

The problems of angiosperm phylogeny are well illustrated by a consideration of the differences between four classifications, all less than a decade old and all by highly respected and experienced authors. The dicotyledons are divided into six subclasses by Cronquist (1981) and seven by Takhtajan (1983), while, for the other two authors, the major groupings are superorders, Thorne (1983) having 19 and Dahlgren (1983) 25. The number of dicotyledonous orders recognized is, respectively, 58, 72, 41, and 83; these figures alone indicate the resulting diversities of names and content, all of which reflect our comparative ignorance of the course that evolution has taken in the angiosperms. In contrast to this, at the next level down the hierarchy, there is basic agreement about the "core" families to be recognized (Heywood, 1978).

Macromolecular sequences provide taxonomic characters whose homology over widely diverse species can be assumed with some confidence. Sequence data can be analyzed objectively with computers. We will probably see in the next decade the publication of nucleic acid sequences long and variable enough to solve some of the problems of

angiosperm phylogeny (e.g., Palmer et al., 1988; Zimmer et al., 1989). It is therefore an appropriate time, when nucleic acid sequencing is supplanting protein sequencing, to set out the results of a decade of work that has produced 335 partial protein sequences from a wide range of angiosperms. These sequences are shorter than nucleic acid sequences already published and therefore contain less information and are less able to resolve the sequential divergences of early radiations. Nevertheless, we believe that our phylogenetic trees will indicate likely relationships and profitable working hypotheses for future investigations.

A SUMMARY OF PUBLISHED INVESTIGATIONS USING PROTEIN SEQUENCES

The pioneer of the use in botany of protein sequences for investigating plant phylogeny was D. Boulter of the University of Durham, England. During the 1970s, Boulter, along with his colleagues and students, published 25 sequences of cytochrome c, 12 complete and 58 partial sequences of plastocyanin and seven sequences of

¹ We acknowledge with thanks grants from the Australian Research Grants Scheme, the Australian Research Council, the Missouri Botanical Garden, the Potter Foundation, the Utah Foundation, and the National Botanic Gardens (Canberra). We are indebted to Richard Norrish who, for many years, has kept our machinery in working order. As indicated in Table 2, we have obtained leaves from many sources, to all of whom we are grateful; we are especially grateful to the Adelaide Botanic Gardens, the Missouri Botanical Garden, the National Botanic Gardens (Canberra), and the Royal Botanic Gardens, Kew.

² Department of Botany, University of Adelaide, Box 498, G.P.O., Adelaide, South Australia 5001.

ferredoxin. These have been collated, with references, by Ramshaw (1982), and Scogin (1981) has reviewed the results from the taxonomic point of view. Although this work generated much interest, it also gave rise to skepticism, some of which can, with hindsight, be attributed to the inadequacies of computing methods that were being developed concurrently. The mostly unfavorable reaction of systematists, epitomized by the review of Cronquist (1976), influenced the cessation of research in Boulter's laboratory about 1980.

Before this, however, partial sequences (up to 25 N-terminal amino acids) of the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (rubisco-SSU) were obtained from six species (Hasslett et al., 1976; Strobaek et al., 1976). This work led to a complete SSU sequence from spinach (Martin, 1979), a forerunner of the work presented here which concerns the N-terminal 40 amino acids of this protein. (The complete sequencing of a protein requires prior purification of several fragments and is at least an order of magnitude more time-consuming than the direct sequencing of the N-terminus of the whole protein using an automatic sequencer.) Nucleotide sequences of rubisco-SSU from a few species have been published, and all of them have been studied using our method. The only new data comparable to our 334 species are from two closely related orchids and their hybrid (G. C. Martin et al., 1987). We are unaware of phylogenetically useful sequences of other proteins since those of Grund et al. (1981) and Nakano et al. (1981).

Work in our laboratory has proceeded in five phases. In phase 1 species were chosen because Boulter had already published their complete sequences of cytochrome c and partial sequences of plastocyanin. When a pattern failed to emerge from analyses of these data, we decided to sample each family with sequences from at least two more representative genera. Thus, the families Apiaceae, Asteraceae, Brassicaceae, Caprifoliaceae, Chenopodiaceae, Fabaceae, Malvaceae, Poaceae, Polygonaceae, Ranunculaceae, and Solanaceae have each been sampled at least three times. These early results were published in a series of papers (Martin et al., 1983; Martin & Dowd, 1984a, b, c).

The sequences for rubisco-SSU, cytochrome c, and plastocyanin were analyzed for these families by Martin, Boulter, and Penny (1985) using derived estimates of familial node sequences. Analyses of data from single macromolecules were not consistent with one another but, for nine of the families, a phylogenetic tree derived from combined data remained consistent when ferredoxin or 5S-

ribosomal RNA (available for some of the families) was added.

This result indicated the need for longer sequences and better sampling of families. Although rubisco-SSU was always multiply represented, in 17 of the 33 samples of other macromolecules there was only a single sequence. This situation is precarious because, if the average distance from a familial node to a species is N , then on the average a single sequence will misrepresent the familial node by N . This source of error might be responsible for part of the poor agreement observed. Sampling a family at least twice, preferably from widely divergent representatives should give a better estimate of the familial node (see phase 5).

In phase 2 we sequenced rubisco-SSU from 11 members of Onagraceae (Martin & Dowd, 1986a), 15 monocotyledons (Martin & Dowd, 1986b), and 14 species of *Solanum* (Martin et al., 1986). We reasoned that the reliability of our methods might be estimated by comparison with taxonomically well understood groups. The results were similar to other taxonomic treatments. Additional species of Asteraceae were also studied and those results will be presented in this paper.

To estimate the rate of evolution, Proteaceae, Solanaceae, Fagaceae, and Winteraceae were sampled in phase 3 using species whose ancestors are thought to have been separated by continental drift at known times. This led to a preliminary publication (Martin & Dowd, 1984b), and the derivation of a molecular evolutionary clock (Martin & Dowd, 1988), which indicated that on average one nucleotide difference arose between two diverging lines once in seven million years.

In phase 4 we tested the hypothesis that leghemoglobin had evolved in plants by lateral transfer from animals. This led to an investigation of all species for which leghemoglobin sequences had been published, and it was shown that the pathway of evolution in those species was closely parallel in hemoglobin and rubisco-SSU (Martin & Dowd, 1986c), suggesting that there was no need to invoke novel evolutionary processes. A consequence of this study was that we increased the number of species of Fabaceae sequenced to eight (see Group 14 below) and obtained sequences from several additional families. Many of these were too small to be studied in the normal course of this investigation but were obtained either because they are known to include nitrogen-fixers or thought to be relatives of the legumes; these include Betulaceae, Casuarinaceae, Chrysobalanaceae, Coriariaceae, Crossosomataceae, Datisceae, Elaeagnaceae, Moringaceae, and Myricaceae.

In phase 5 we surveyed the dicotyledons which increased the number of families studied from 24 to 124.

A SURVEY OF THE DICOTYLEDONS

There are about 250 families of dicots. Because it was impractical to sample all of them, a decision was made to sample about half, i.e., to increase the number from the 24 mentioned above to 124. Three families (Acanthaceae, Loranthaceae, Santalaceae) failed for reasons that will be discussed later. The additional 97 families were chosen primarily on the basis of size. The majority of families sampled have more than 20 genera. To cover as wide a range of variation as possible, some small families were also sampled. For example, the order Illiciales has only three genera, so the family Schisandraceae (two genera) was chosen to represent it. Only three orders are unrepresented out of Thorne's 41 (two of which are parasitic and devoid of rubisco), 10 out of Cronquist's 58, 19 out of Takhtajan's 72, and 21 out of Dahlgren's 85.

It is impractical, mainly because computers are limited in their capacities to analyze large numbers of taxa simultaneously, to contemplate building a phylogenetic tree for 122 families (comprising 310 species) without some subdivision into groups. We have done this by referring to all four current phylogenies. Thorne (1983) and Dahlgren (1983) have superorders as their major groups, the former nominating 19 and the latter 25. If these two authors agree that families are in the same superorder then they have been grouped together in our scheme, with one proviso. Takhtajan (1983) and Cronquist (1981) have respectively seven and six subclasses as their major groups, and these two authors have been allowed a veto; if either if them does not also agree that families are in the same subfamily, then they are left ungrouped. In this way we have divided 102 of the studied families into 25 Groups, leaving 20 ungrouped because there is disagreement. We are reluctant to use a formal term like superorder but need to make it clear that our use of Group does have a defined meaning, so we have used a capital G. The Groups are shown in Table 1.

It was practicable to sample each new family only twice, and we have done this by choosing two species not only from different genera but, if possible, from different subfamilies or tribes. Sometimes this criterion has broken down because fresh leaves have not been available.

In Table 2 the 335 species for which sequences

are available are arranged by families and Groups, and their sources and sequences are given.

BIOCHEMICAL METHODS

The methods published by Martin and Jennings (1983) have stood the test of time, so, rather than repeat them here, a general description will be given and the few modifications mentioned.

Two methods were described, one for "pungent" leaves with high concentrations of phenolics or other substances that make protein purification difficult, the other for "bland" species whose leaves are much more amenable. The bland method gives better quality protein and is therefore to be preferred. However, because the pungent method works well with bland leaves, but not vice versa, it was preferred when there was doubt or too few leaves for trial extractions.

Both procedures started with maceration of about 100 g of leaves from which the midribs were removed if practicable. For bland leaves the extracting buffer was essentially a reducing, saline tris-HCl buffer at pH 7.4, while for pungent leaves a reducing, saline borate buffer at pH 8.6 and containing the detergent Triton X-100 was used. After crude straining and centrifugation to remove solids, the extract was passed through a succession of two liquid gel columns. A Sephadex G-25 column was used first to remove low molecular weight substances. A Sepharose 6B column was used to remove remaining low molecular weight substances and high molecular weight nucleic acids and membrane fragments. Eluting buffers were different for the two extraction procedures and for the different columns used. The protein was precipitated with ammonium sulfate for the bland method and with acetone for pungent. Procedures after the second column were the same for both types of leaves. The protein was S-carboxymethylated at pH 8.6 to break disulphide bridges between cysteine residues and then passed through a long column of Sephadex G-100 in an eluting buffer containing sodium dodecyl sulfate. This separated the large subunit from the small subunit, which was precipitated in acetone and dried before sequencing. (A variation of this procedure was to use a column of G75 followed by G-100.)

The methods are rather crude but are successful because rubisco is a very large protein and, by a considerable margin, the most abundant protein in leaves.

About 5 mg of small subunit (in 0.5 ml of water without polybrene) was sequenced on the Beckman

TABLE 1. Families of dicotyledons grouped because they are placed in the same major taxon by all of Cronquist (1981), Dahlgren (1983), Takhtajan (1983), and Thorne (1983).

GROUP 1	GROUP 4	GROUP 9	GROUP 12	GROUP 17	GROUP 21
Magnoli	Ulm	Dipterocarp	Eric	Connar	Lami
Winter	Mor	Elaeocarp	Epacrid	Sapind	Verben
Annon	Urtic	Tili	GROUP 13	Anacardi	GROUP 22
Myristic	GROUP 5	Sterculi	Cunoni	Simaroub	Solan
Schisandr	Hamamelid	Bombac	Ros	Meli	Convolvul
Monimi	Betul	Malv	Saxifrag	Rut	Polemoni
Laur	Fag	GROUP 10	GROUP 14	GROUP 18	GROUP 23
Aristoloch	Casuarin	Viol	Caesalpini	Halorag	Scrophulari
Calycanth	GROUP 6	Flacourti	Mimos	Rhizophor	Gesneri
GROUP 2	Dilleni	Datisc	Papilioni	GROUP 19	Bignoni
Berberid	Thea	Cucurbit	GROUP 15	Zygophyll	Pedali
Ranuncul	Ochn	Salic	Trap	Gerani	GROUP 24
Lardizabal	Clusi	Cappar	Lyth	Tropaeoli	Valerian
Menisperm	GROUP 7	Brassic	Myrt	Malpighi	Caprifol
Papaver	Myric	Resed	Punic	GROUP 20	GROUP 25
GROUP 3	Jugland	Moring	Onagr	Logani	Api
Cabomb	GROUP 8	GROUP 11	Melastomat	Gentian	Arali
Nymphae	Caryophyll	Sapot	Combret	Apocyn	
	Nyctagin	Styrac	GROUP 16	Asclepiad	
	Amaranth	Primul	Olac	Ole	
	Phytolacc	Myrsin	Celastr	Rubi	
	Chenopodi				
FAMILIES THAT DO NOT FIT INTO ONE OF THE GROUPS					
Aster	Coriari	Goodeni	Nelumbon	Polygon	Thymelae
Bux	Crossosomat	Hydrophyll	Piper	Prote	Vit
Campanul	Elaeagn	Lecythid	Plumbagin	Rhamn	
Chrysobalan	Euphorbi	Loas			

Note: "-aceae" omitted from all names.

890C automatic sequencer using Beckman's standard quadrol program with 50% quadrol buffer. The phenylthiohydantoin (PTH) derivatives of the amino acids were identified using a Waters HPLC instrument with a C-18 radially compressed column and eluted with 0.1M sodium acetate (pH 6.0) and acetonitrile. This did not distinguish two pairs of amino acids and was therefore supplemented with TLC.

Using these methods, we could, without assistance, produce two proteins each week and sequence two others.

FAILURES

Although 90% of attempts led to successful sequences, the remaining 10% deserve brief attention. Unless there was an identified reason for failure that could be corrected, our policy was to try another representative of the family.

Faults that could be corrected include the amounts of extraction and elution buffers used. Some plants gave extracts that were mucilaginous to the point of setting solid. Dilution of the extract

corrected this. This problem occurred in Onagraceae and a few others with small leaves containing a high proportion of veins. Insolubility of the protein, leading to precipitation in columns, could sometimes be corrected by loading a more dilute extract. Plants with C4 photosynthesis, and rubisco tightly bound in bundle sheaths, were avoided if possible. Plants with C3 photosynthesis often occur in the same genera or families and were unlikely to be phylogenetically biased. However, if unavoidable (e.g., *Welwitschia* is reported to be C4), special care was taken during the maceration process.

It is suspected that the most common cause of failure was the presence of powerful proteases in the leaves and, in retrospect, it would have been profitable to try correcting this with research early in the project. Species of *Ficus*, known to have leaf proteases, showed symptoms of this failure. Large amounts of protein traveled where the small subunit should have been on the G-100 column and gave many amino acids at each position when sequenced. Another casualty of this sort was *Gnetum gnemon*, which was particularly desired because it is a gymnosperm thought to be close to

TABLE 2. Species studied and sequences.

Notes:	
(1) The dicotyledons are arranged by Groups as explained in the text and Table 1. After each family name we give the three-letter abbreviation of Weber(1982).	
(2) After the species name and author, a slash is followed by the source of the specimen: P.G.=Private Garden (in Adelaide unless otherwise stated); B.G.=Botanic(al) Garden(s). Australian states: N.S.W.=New South Wales; N.T.=Northern Territory; Qld=Queensland; S.A.=South Australia; Vic.=Victoria; W.A.=Western Australia. N.Z.=New Zealand. C.S.I.R.O.=Commonwealth Scientific and Industrial Research Organisation. C.S.=certified seed.	
(3) Key to N-terminal 40 amino acids: A, alanine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine.	
D I C O T Y L E D O N S.	
GROUP 1	
ANNONACEAE ANN	
Desmos goezeanus (F.Muell.)Jessop/N.Qld	M M V W P P E G K K K F E T L S Y L P S L S S E Q L A K E V E Y L L L N G W I P
Xylopia maccraei (F.Muell.)L.S.Smith/N.Qld	M M V W P P E G K K K F E T L S Y L P T L S S E E L A K E V E Y L I I S G W I P
ARISTOLOCHIACEAE ARS	
Aristolochia baetica L./Adelaide B.G.	M Q V W P P E G L K K F E T L S Y L P T L T E E Q L A K E I E Y L L L K G W V P
Asarum canadense L./Missouri B.G.	M Q V W P P E G L K K F E T L S Y L P P L S S Q E L A K E V D Y L L I K G W I P
CALYCANTHACEAE CAL	
Calycanthus floridus L./Adelaide B.G.	M Q V W P P E G L K K F E T L S Y L P P L S T E Q L I K E V D Y L L L K G W V P
Chimonanthus praecox (L.)Link./Adelaide B.G.	M Q V W P P E G L K K Y E T L S Y L P P L S S E Q L I K E I D Y L L L V G W V P
Idiospermum australiense (Diels)S.T.Blake/N.Qld	M Q V W P P V/EG L P K F E T L S Y L P P L V T E Q I/LA K E I D Y L L L N G W V P
LAURACEAE LAU	
Laurus nobilis L./Adelaide B.G.	M L V W P P/TV/EG L K K F E T L S Y L P P L S S Q Q L A K E V D Y L L I S G W V P
Persea americana Mill./P.G.	M M V W P P V/EG L K K F E T L S Y L P P L S P E Q L A A E V D Y L L S G W V P
MAGNOLIACEAE MAG	
Liriodendron tulipiferum L./P.G.	M L V W P P V G L K K F E T L S Y L P P L S P Q S L A K E V D Y L L R S G W I P
Magnolia soulangiana Soul.-Bod./P.G.	M L V W P P E G L K K F E T L S Y L P P L S S E E L A K E I D Y L I G S G W I P
Michelia figo (Lour.)Sprengl/Adelaide B.G.	M L V W P P E G L K K F E T L S Y L P P L S S E E L A K E I D Y L I R S G W I P
MONIMIACEAE MNM	
Peumus boldus Molina/Adelaide B.G.	M M V W P P E G L K K F E T L S Y L P P L S S E A L A K E V D Y L L I M G W V P
MYRISTICACEAE MYS	
Knema latericia Elmer/Singapore B.G	M M V W P P L G N K K F E T L S Y L P P V S V/ME Q L A K E I D Y L L L N G W V P
Myristica fragrans Houtt./Singapore B.G.	M M V W P P L G S K K F E T L S Y L P P L S L E Q L A K E I D Y L L L N G W V P
SCHISANDRACEAE SCS	
Kadsura japonica Dunal/Berkeley B.G.	M K V W P P Y N M K K F E T L S Y L S P I S Q E E L I K E I E Y L L L K G W V P
Schisandra rubriflora Rehder & Wilson/Kew B.G.	M Q V W P P Y N M K K F E T L S Y L S P L S Q E E L L K E I E Y L L L N G W V P
WINTERACEAE WIN	
Drimys winteri Forster/Kew B.G.	M K V W P P K G L L K F E T L S Y L P P L S S E Q L A K Q V E Y L I K M G W V P
Pseudowintera colorata (Raoul)Dandy/North N.Z.	M K V W P P K G L L K F E T L S Y L P P L S P Q Q L A K Q I D Y L I Q S G W V P
Tasmannia purpurascens (J.W.Vickery)A.C.Smith/ Canberra B.G.	M K V W P P K G L L K F E T L S Y L P P L S P G Q L A K Q V D Y L I K S G W V P
GROUP 2	
BERBERIDACEAE BER	
Berberis aetnensis Presl/Adelaide B.G.	M Q V W P P V G L K K F E T L S Y L P P L S S Q S L A K E V D Y L L L V G W V P
Nandina domestica Thunb./P.G.	M Q V W P P I G L K K F E T L S Y L P P L S P E A L A K E V D Y L L L K G W V P

TABLE 2. Continued.

Plant Name	Location	Accession Number
LARDIZABALACEAE LAR		
Akebia quinata (Houtt.)Decne./Adelaide B.G.		M Q V W P P I G L K K F E T L S Y L P P L T Q E Q L A K E V E Y L L S S G W V P
Decaisnea fargesii Franchet/Kew B.G.		M Q V W P P V N L K K F E T L S Y L P P L S Q E Q L L K E V D Y L L I N N G W I P
Lardizabala biternata Ruiz & Pav./Edinburgh B.G.		M K V W P P V G L K K F E T L S Y L P P L S Q E A L A K E I D Y L L S N G W V P
MENISPERMACEAE MNS		
Cocculus trilobus (Thunb.)DC./Adelaide B.G.		M K V W P P L G L K K F E T L S Y L P P L S P Q A L A K E V D Y L L L N G W I P
Hypserpa laurina (F.Muell)Diels/Canberra B.G.		M K V W P P V L G L K K F E T L S Y L P P L S P Q A L A K E V D Y L L L K G W I P
PAPAVERACEAE PAP		
Eschscholtzia californica Cham./Adelaide B.G.		M Q V W P P T G L K K F E T L S Y L P P L T A E Q L A K E V D Y L L K N G W V P
Papaver orientale L./P.G.		M Q V W P P T G L K K F E T L S Y L P P D L T V E Q L A K E I D Y L L R N G W V P
Dicentra spectabilis Lem./P.G.		M Q V W P P Y N N K K F E T L S Y L P P D L T V E Q L A K E V D Y L L L K G W V P
Fumaria officinalis L./P.G.		M K V W P P Y G M K K F E T L S Y L P Q L S Q E E L A K E V D Y L L I S N G W V P
RANUNCULACEAE RAN		
Aquilegia vulgaris L./P.G.		M K V W P P L G L L K K F E T L S Y L P P L S Q Q A L A K E V D Y L L A N G W V P
Clematis rehderiana Craib/Adelaide B.G.		M K V W P P Y N N K K F E T L S Y L P P E L T P E S L A K E I E / D Y L L I R K G W V P
Clematis viticella L./Adelaide B.G.		M K V W P P Y N N K K F E T L S Y L P P E L T P E S Q I L K E I D Y L L I R K G W V P
Helleborus foetidus L./Adelaide B.G.		M K V W P P L N N L K K F E T L S Y L P P L S Q A S L A K E V D Y L L I V K G W I P
Ranunculus trilobus Desf./P.G.		M K V W N P V G N Q K F E T L S Y L P P L S T E S L A K E V E Y L L I R K G W V P
GROUP 3		
CABOMBACEAE CAB		
Brasenia schreberi Gmelin/Nagambie Vic.		M A V W P I Y N / V L K K F E T L S Y L P S L S K E Q L A K E V E Y L L I K K W I P
NYMPHAEACEAE NYM		
Nuphar advena W.T.Aiton/Canberra Public garden		M K V W P P Y N / D V Q / K K Y E T L S Y L P P L S T E Q L A K E V D Y L L S N K W I P
Nymphaea gigantea Hook./P.G.		M K V W N P Y N N I K F E T L S Y L P P L T D E Q L A K E V E Y L L L S K W I P
Victoria amazonica (Poeppig)Sowerby/Adelaide B.G.		M K V W P P Y N N I K F E T L S Y L P P L S D E Q L A K E V E Y L L L K K W I P
GROUP 4		
MORACEAE MOR		
Humulus scandens (Lour.)Merr./Adelaide B.G.		M K V W P P S N V K K F E T L S Y L P P L S V E Q L L K E V E Y L L K S G W V P
Morus alba L./Adelaide B.G.		M K V W P P S G L L K K F E T L S Y L P P D L T S E Q L I K E I D Y L L L S N W I P
ULMACEAE ULM		
Parasponia andersoni Planch./CSIRO Canberra		M Q V W P P L G K K K F E T L S Y L P P E L T I E Q L I K Q I E F L I K K G W V P
Ulmus parvifolia Jacq./Adelaide B.G.		M Q V W P P P G Y K K F E T L S Y L P P D L T T E Q L I K E V E Y L L L K G W V P
URTICACEAE URT		
Parietaria judaica L./Adelaide B.G.		M Q V W P P S G L K K F E T L S Y L P P D L T E E Q L I K E V D Y L L I K S K L / W I / V P
Pipturus argenteus Wedd./Adelaide B.G.		M Q V W P P S G L K K F E T L S Y L P P L T E A Q L L K Q I D F L I K S K L I P
GROUP 5		
BETULACEAE BET		
Alnus maximowiczii Callier/Adelaide B.G.		M K V W P P L G L K K K F E T L S Y L P P L T A E Q L A K E V D Y L L L K G W I P
Betula atrata Domin./Adelaide B.G.		M Q V W P P L G L K K K F E T L S Y L P P L T T E Q L A K Q V D Y L L L K N G W I P
CASUARINACEAE CAS		
Allocasuarina muelleriana (Miq.)L.Johnson/ Adelaide B.G.		M Q V W P P L G K K K K F E T L S Y L P P L S S E S L A K E V D Y L L L K K W I P
Casuarina glauca Sieber/Adelaide B.G.		M Q V W P P L G K K K K F E T L S Y L P N L S S E S L A K E V D Y L L L K K W I P
Gymnostoma (sp. nov. L.Johnson)/N.Qld		M Q V W P P S G K K K F E T L S Y L P Q L S S E S L A K E V D Y L L L K K W I P
FAGACEAE FAG		
Fagus sylvaticus L./P.G.		M K V W P P L G L Q K K F E T L S Y L P P L V S E N L A K Q I E Y L L N K G W I P
Nothofagus aequilateralis (Baum.-Bodenh.)Steenis/ Adelaide B.G.		M K V W P P L G L I K K F E T L S Y L P P L S S E I L A K E V D Y L L I S S G W I P

TABLE 2. Continued.

New Caledonia	
Nothofagus antarctica (Forster)Oersted/Kew B.G.	M K V W P P L G L K K F E T L S Y L P T L T P A E L A K E V D Y L L R K G W V P
Nothofagus betuloides (Mirbel)Oersted/Kew B.G.	M K V W P P L G L K K F E T L S Y L P T L T P E N L A K E V D Y L L N K G W V P
Nothofagus cunninghamii Ostenfeld/Cape Otway Vic.	M K V W P P L G L K K F E T L S Y L P T L T N E N L A K E V D Y L L R K G W V P
Nothofagus fusca (Hook.f.)Oersted/Otari B.G. N.Z.	M K V W P P L G L K K F E T L S Y L P T L S S E A L A K E V D Y L I S K G W V P
Nothofagus gunnii Hook.f./Tasmania	M K V W P P L G L K K F E T L S Y L P T L S P E Q L A L K V D Y L L R K G W V P
Nothofagus menziesii (Hook.f.)Oersted/ Wellington B.G. N.Z.	M K V W P P L G L K K F E T L S Y L P T L S S E S L A K E V D Y L L R K S W V P
Nothofagus moorei F.Muell/New England NSW	M K V W P P L G L K K F E T L S Y L P T L T P E S L A K E V D Y L L R K G W V P
Nothofagus procera (Poeppig & Endler)Oersted/ Kew B.G.	M K V W P P L G L K K F E T L S Y L P T L T P A E L A K E V D Y L L R K G W V P
Nothofagus pumilio (Poeppig & Endler)Krasser/ Kew B.G.	M K V W P P L G L K K F E T L S Y L P T L S P E S L A K E V D Y L L L K G W V P
Nothofagus solandri (Hook.f)Oersted/Otari B.G. N.Z.	M K V W P P L G L K K F E T L S Y L P T L S L E A L A K E V D Y L I S K G W V P
Quercus robur L./P.G.	M Q V W P P L G L K K Y E T L S Y L P T L T D E Q L G R E I D Y L L R S G W V P
HAMAMELIDACEAE HAM	
Liquidambar styraciflua L./Adelaide B.G.	M K V W P P L G L K K F E T L S Y L P T L S N A A L A K E V D Y L L R S G W V P
Parrotia persica C.Meyer/Adelaide B.G.	M K V W P T E N N L K F E T L S Y L P T L T N A D L A K E V D Y L L R S G W I P
GROUP 6	
CLUSIACEAE CLU	
Garcinia mangostana L./CSIRO Darwin N.T.	M Q V W P P L G L K K F E T L S Y L S P L S L A E L A K E V D Y L L L S K W I P
Hypericum hookerianum Wight & Arn./Adelaide B.G.	M Q V W P P L G L K K F E T L S Y L P P L S I E Q L A K E V D Y L I I S G W I P
DILLENIACEAE DLL	
Dillenia alata (DC.)Martelli/Canberra B.G.	M K V W P P S G L L K K F E T L S Y L P D L S S Q E L A K E V D Y L L R M G W I P
Hibbertia scandens (Willd.)Gilg/P.G.	M L V W P P E G L K K F E T L S Y L P P L S P Q E L L K / L E V A Y L L L K G L I P
OCHNACEAE OCH	
Brackenridgea nitida A.Gray/N.Qld	MK/QV W P T E G L S K F E T L S Y L P D L T E A Q L I K E I D Y L I I K G W I P
Ochna atropurpurea DC./Adelaide B.G.	M K V W P T E G L K K F E T L S Y L P S L T Q A Q L I K E V D Y L L I K G W I P
THEACEAE TEA	
Gordonia axillaris Szyszyl./Mt Lofty B.G. S.A.	M K V W S P L G L K K F E T L S Y L P S L T N I Q L A K E V D Y L I N K G W I P
Ternstroemia cherryi (F.M.Bailey)Merr. ex J.F.Bailey& C.T.White/N.Qld	M M V W P P L G L K K F E T L S Y L P S L S D I E L A K E V D Y L L L S G W V P
GROUP 7	
JUGLANDACEAE JUG	
Carya illinoensis (Wangenh.)Koch/Adelaide B.G.	M Q V W P P L G K K K F E T L S Y L P P L S P A Q L A K E V D Y L L S K G W I P
Juglans regia L./Adelaide B.G.	M Q V W P P L G K K K F E T L S Y L P P L T A E S L A K E V D Y L I S K G W I P
MYRICACEAE MYR	
Myrica pennsylvannica Loisel/Adelaide B.G.	M Q V W P P L G K K K F E T L S Y L P P L S P A A L A K E V D Y L L R K G W I P
GROUP 8	
AMARANTHACEAE AMA	
Bosea amherstiana (Moq.)Hook./Adelaide B.G.	M Q V W P P L G K / L K K F E T L S Y L P P L S Q E S L T K E V S Y L L L K G W V P
Celosia argentea L./P.G.	M Q V W P P T / L G K K K F E T L S Y L S P L S Q E Q L L K Q V S Y L L N K G W V P
CARYOPHYLLACEAE CRY	
Cerastium tomentosum L./Adelaide B.G.	M Q V W P P T G L K K K F E T L S Y L P P L S E E Q L L K E I N Y L L I K G W V P
Silene colorata Poirret/Adelaide B.G.	M Q V W P P S N L K K F E T L S Y L P P L T E E Q L L K E I N Y L L V K G W V P
CHENOPODIACEAE CHN	
Beta vulgaris L./C.S.	M Q V W P P L G L K K K F E T L S Y L P P L T T E Q L L A E V N Y L L V K G W I P
Chenopodium murale L./P.G.	M Q V W P P V G K K K F E T L S Y L P P L S P E S L L G E V E Y L L N K G W V P

TABLE 2. Continued.

Spinacia oleracea L./C.S.	M Q V W P P L G L K K F E T L S Y L P P L T T E Q L L A E V N Y L L V K G W I P
NYCTAGINACEAE NYC	
Bougainvillea glabra Choisy/Adelaide B.G.	M Q V W P P T G K I K F E T L S Y L P P L T P E Q L L K E V Q Y L L N K G W V P
Mirabilis jalapa L./Adelaide B.G.	M Q V W P P I/LG K K I F E T L S Y L P T L T S E Q L L K E V Q Y L L N K G W V P
PHYTOLACCACEAE PHT	
Codonocarpus cotinifolius (Desf.)F.Muell./North S.A	M Q V W P T V G L K K F E T L S Y L P D L N E E S L A K E V D Y L L L K G W I P
Phytolacca dioica L./Adelaide B.G.	M Q V W P T V/LGL/KK K F E T L S Y L P P L S Q E S L T K E V E Y L L N K G W V P
GROUP 9	
BOMBACACEAE BOM	
Chorisia insignis Kunth/Adelaide B.G.	M Q V W P P L G K K K F E T F S Y L P D L V T E E L I K E I E Y L I L K K W I P
Durio zibethinus L./CSIRO Darwin N.T.	M K V W P P L G I K K F E T F S Y L P D L S T D E L L K E I E Y L L L K K W I P
DIPTEROCARPACEAE DPC	
Dryobalanops aromatica Gaertner/Singapore B.G.	M Q V W P P L G K K K F E T L S Y L P D L T P E Q L A K E V D Y L L L S K W I P
Shorea pauciflora King/Singapore B.G.	M Q V W P P L G K K K F E T L S Y L P D L S P E Q L A K E V D Y L L L S K W V P
ELAEOCARPACEAE ELC	
Aristotelia australasica F.Muell./Canberra B.G.	M Q V W P P L G K K K F E T L S Y L P P L S A A E L A K E V D Y L L L K G W V P
Eleocarpus grandis F. Muell./Adelaide B.G.	M Q V W P P I G L K K F E T L S Y L P P L S T A A L A K E I D Y L L L K G W V P
MALVACEAE MLV	
Althaea rosea (L.)Cav./P.G.	M K V W P P L G L K K F E T L S Y L P P L T T E Q L A K E V E Y L L R K K W V P
Malva parviflora L./P.G.	M K V W P P L G L K K F E T L S Y L P P L T S E Q L A K E V E Y L L R K K W V P
Sphaeralcea umbellata (Cav.)G.Don/Adelaide B.G.	M K V W P P L G L K K F E T L S Y L P T L T D E Q L A K E V E Y L L R S K W V P
STERCULIACEAE STR	
Brachychiton discolor F.Muell./Adelaide B.G.	M K V W P P L G L K K F E T L S Y L P P L S T E E L L K E V E Y L I L S N W I P
Dombeya mastersi Hook f./Adelaide B.G.	M K V W P P E G L K K F E T L S Y L P D L S T E Q L L K E I D Y L I I S N W I P
TILIACEAE TIL	
Grewia occidentalis L./Adelaide B.G.	M K V W P P I N K K I F E T L S Y L P P L T V T Q L A K E V D Y L L L N G W I P
Sparmannia africana L.f./Adelaide B.G.	M K V W P T E G L K K F E T L S Y L P T L T D E Q L L K E I D Y L I R K K W I P
GROUP 10	
BRASSICACEAE BRA	
Brassica napus L./C.S.	M K V W P P V G/AK K K K F E T L S Y L P D L T E V E L G K E V D Y L L E N K W I P
Nasturtium officinale R.Br./P.G.	M Q V W P P I G K K K F E T L S Y L P D L T D V E L G K E V D Y L L R N K W I P
Raphanus sativus L./C.S.	M K V W P P I G K K K F E T L S Y L P D L T D V E L G K E V D Y L L R N K W I P
CAPPARACEAE CPP	
Cleome spinosa L./P.G.	M Q V W P T T G L R K F E T L S Y L P D L S D D E L A K E V D Y L L N N W I P
CUCURBITACEAE CUC	
Cucumis sativus L./P.G.	M K V W P P L G L R K F E T L S Y L P D M S N E Q L S K E R D Y L L S N G W V P
Cucurbita pepo L./P.G.	M K V W P P L G L R K F E T L S Y L P D L S D A Q L L K E I G Y L L L S G W V P
DATISCEAE DAT	
Datisca cannabina L./Adelaide B.G.	M K V W P P L G L K K Y E T L S Y L P P/DL T T E Q L S K E V E Y L I I V G W I P
Tetrameles nudiflora R.Br./N.Qld	M Q V W P P V G K K K F E T L S Y L P P L T D E Q L S K E V E Y L L I S K W I P
FLACOURTIACEAE FLC	
Azara celastrina D.Don/Adelaide B.G.	M Q V W P P V G L K K F E T L S Y L P D L T S E Q L A K E I E Y L L I S K W V P
Kiggelaria africana L./Adelaide B.G.	M Q V W P P L G K K K F E T L S Y L P D L T S E E L A K E V E Y I L L N K W I P
MORINGACEAE MRG	
Moringa oleifera Lam./Lyon Arboretum Hawaii	M Q V W P P T G L K K F E T L S Y L P P L S A E E L A K E V D Y L L L K G W I P
RESEDACEAE RSD	
Reseda luteola L./Adelaide B.G.	M Q V W P P L G L K K F E T L S Y L P D L S D E E L A K E V D Y L L S N K W I P
SALICACEAE SAL	

TABLE 2. Continued.

Populus nigra L./Adelaide B.G.	M	Q	V	W	P	P	L	G	L	K	K	F	E	T	L	S	Y	L	P	D	L	S	T	E	Q	L	A	K	E	I	Q	Y	L	L	L	N	K	W	V	P	
Salix babylonica L./Adelaide B.G.	M	Q	V	W	P	P	V	L	G	L	K	K	F	E	T	L	S	Y	L	P	D	L	T	K	E	Q	L	A	K	E	I	E	Y	L	L	L	S	K	W	V	P
VIOLACEAE VIO																																									
Melicytus ramiflorus Forster/Adelaide B.G.	MK/QV	W	P	T	V	G	L	K	K	F	E	T	L	S	Y	L	P	P	L	T	S	E	Q	L	A	K	E	I	D	Y	L	I	I	K	G	W	I	P			
Viola odorata L./P.G.	M	K	V	W	P	T	T	G	L	K	K	F	E	T	L	S	Y	L	P	P	L	T	T	E	Q	L	A	K	E	I	D	Y	L	I	S	N	G	W	I	P	
GROUP 11																																									
MYRSINACEAE MRS																																									
Ardisia humilis Vahl/Adelaide B.G.	M	K	V	W	P	P	L	G	L	K	K	F	E	T	L	S	Y	L	P	P	L	T	E	A	Q	L	A	K	E	I	E	Y	L	L	L	S	G	W	V	P	
Rapanea variabilis (R.Br.)Mez/Adelaide B.G.	M	K	V	W	P	P	L	G	L	K	K	F	E	T	L	S	Y	L	P	P	L	T	I	E	Q	L	A	K	E	I	D	Y	L	L	L	S	G	W	I	P	
PRIMULACEAE PRM																																									
Cyclamen neapolitanum Ten./P.G.	M	K	V	W	P	P	L	G	L	K	K	F	E	T	L	S	Y	L	P	P	L	S	S	E	Q	L	L	K	E	V	D	Y	L	L	L	N	G	W	V	P	
Primula malacoides Franchet/P.G.	M	K	V	W	P	P	L	G	L	K	K	F	E	T	L	S	Y	L	P	P	L	T	V	E	Q	L	L	K	E	V	D	Y	I	I	I	N	G	W	I	P	
SAPOTACEAE SPT																																									
Mimusops cyanocarpa Cunn. ex Sweet/Adelaide B.G.	M	T	V	W	P	P	L	G	L	K	V	F	E	T	L	S	Y	L	P	P	L	T	Y	E	E	L	A	A	E	V	D	Y	L	I	I	S	K	W	I	P	
Planchonella australis (F.Muell.)Pierre/ Adelaide B.G.	M	S	V	W	P	P	K	G	L	K	L	F	E	T	L	S	Y	L	P	P	L	S	Y	E	E	L	A	A	E	V	D	Y	I	L	L	S	K	W	I	P	
STYRACACEAE STY																																									
Halesia carolina Ellis/Missouri B.G.	M	K	V	W	P	P	L	G	L	K	K	F	E	T	L	S	Y	L	P	P	L	S	D	N	E	L	L	K	E	I	N	Y	L	I	L	S	G	W	I	P	
Styrax japonica Sieber & Zucc./Missouri B.G.	M	M	V	W	P	T	E	G	L	K	K	F	E	T	L	S	Y	L	P	P	L	T	D	E	E	L	L	K	E	I	E	Y	I	/L	L	L	S	G	W	I	P
GROUP 12																																									
EPACRIDACEAE EPC																																									
Astroloma conostephioides (Sonder)F.Muell./ Montacute S.A.	M	K	V	W	P	P	I	G	I	K	K	F	E	T	L	S	Y	L	P	P	L	T	D	E	Q	L	L	K	E	V	D	Y	L	I	N	K	G	W	I	P	
Leucopogon parviflorus Lindley/Sellick's Beach S.A.	M	K	V	W	P	P	I	G	I	K	K	F	E	T	L	S	Y	L	P	P	L	T	D	Q	Q	L	L	K	E	V	D	Y	L	I	I	K	G	W	V	P	
ERICACEAE ERI																																									
Arbutus andrachne L./Adelaide B.G.	M	K	V	W	P	P	L	G	L	K	K	F	E	T	L	S	Y	L	P	P	L	S	I	P	S	L	A	K	E	I	E	Y	L	I	I	Q	G	W	I	P	
Rhododendron ponticum L./Mt Lofty B.G. S.A.	M	K	V	W	P	P	I	G	L	V	K	F	E	T	L	S	Y	L	P	P	L	S	T	E	Q	L	L	K	E	V	D	Y	L	I	I	K	G	W	V	P	
GROUP 13																																									
CUNONIACEAE CUN																																									
Aphanopetalum resinosum Endl./Canberra B.G.	M	K	V	W	P	P	V	G	L	K	Y	F	E	T	L	S	Y	L	P	P	L	S	Q	A	Q	L	A	K	E	V	D	Y	L	L	L	K	G	W	V	P	
Callicoma serratifolia Andr./Canberra B.G.	M	K	V	W	P	P	L	G	L	K	K	F	E	T	L	S	Y	L	P	P	L	T	T	A	E	L	A	K	E	V	D	Y	L	L	L	S	G	W	I	P	
ROSACEAE ROS																																									
Prunus ilicifolia Walp./Adelaide B.G.	M	K	V	W	P	P	L	G	L	K	K	F	E	T	L	S	Y	L	P	P	L	S	P	E	E	L	A	K	E	V	D	Y	L	L	L	K	N	W	V	P	
Pyrus communis L./Adelaide B.G.	M	Q	V	W	P	P	L	G	L	K	K	F	E	T	L	S	Y	L	P	P	L	S	T	E	E	L	A	K	E	V	D	Y	L	L	L	K	N	W	V	P	
SAXIFRAGACEAE SAX																																									
Bergenia X schmidtii Silva-Tarouca/Adelaide B.G.	M	K	V	W	P	P	L	G	L	K	K	F	E	T	L	S	Y	L	P	P	L	T	Q	E	E	L	A	K	E	V	E	Y	L	L	L	S	G	W	V	P	
Hydrangea paniculata Sieber/P.G.	M	K	V	W	P	T	A	V	L	K	K	F	E	T	L	S	Y	L	P	P	L	T	V	E	Q	L	G	K	E	V	D	Y	L	L	L	K	/R	K	W	V	P
GROUP 14																																									
CAESALPINIACEAE CSL																																									
Ceratonia siliqua L./Adelaide B.G.	M	Q	V	W	P	P	L	G	L	K	K	F	E	T	L	S	Y	L	P	P	L	S	S	E	E	L	G	K	E	I	D	Y	L	L	L	K	G	W	I	P	
Gleditsia triacanthos L./Adelaide B.G.	M	K	V	W	P	P	L	G	L	K	K	F	E	T	L	S	Y	L	P	P	L	S	T	E	E	L	A	K	E	V	D	Y	L	L	R	K	G	W	I	P	
MIMOSACEAE MIM																																									
Acacia fimbriata Cunn. ex G.Don/Adelaide B.G.	M	L	V	W	P	P	L	G	K	K	K	F	E	T	L	S	Y	L	P	P	L	T	T	E	Q	L	A	K	Q	V	D	Y	L	L	L	Q	G	W	V	P	
Leucaena glauca Benth./Adelaide B.G.	M	L	V	W	P	P	L	G	K	K	K	F	E	T	L	S	Y	L	P	T	L	T	T	E	E	L	L	A	E	V	E	Y	L	L	R	Q	G	W	V	P	
PAPILIONACEAE PPL																																									
Glycine max Merr./P.G.	M	Q	V	W	P	P	I	G	K	K	K	F	E	T	L	S	Y	L	P	D	L	D	A	Q	L	A	K	E	V	E	Y	L	L	R	K	G	W	I	P		

TABLE 2. Continued.

Plant Name	Collector	Plant Name	Collector
Goodia lotifolia Salisb./Canberra B.G.		Goodia lotifolia Salisb./Canberra B.G.	
Hovea acutifolia A.Cunn./Canberra B.G.		Hovea acutifolia A.Cunn./Canberra B.G.	
Lupinus angustifolius L./C.S.		Lupinus angustifolius L./C.S.	
Oxylobium lanceolatum (Vent.) Druce/Canberra B.G.		Oxylobium lanceolatum (Vent.) Druce/Canberra B.G.	
Pisum sativum L./C.S.		Pisum sativum L./C.S.	
Vicia faba L./C.S.		Vicia faba L./C.S.	
Vigna radiata (L.) Wilezek/C.S.		Vigna radiata (L.) Wilezek/C.S.	
GROUP 15		GROUP 15	
COMBRETACEAE		COMBRETACEAE	
Combretum decandrum Roxb./Adelaide B.G.		Combretum decandrum Roxb./Adelaide B.G.	
Quisqualis indica L./Missouri B.G.		Quisqualis indica L./Missouri B.G.	
LYTHRACEAE		LYTHRACEAE	
Lythrum salicaria L./Adelaide B.G.		Lythrum salicaria L./Adelaide B.G.	
Woodfordia fruticosa (L.) Kurz/Adelaide B.G.		Woodfordia fruticosa (L.) Kurz/Adelaide B.G.	
MELASTOMACEAE		MELASTOMACEAE	
Melastoma affine D. Don/North N.S.W.		Melastoma affine D. Don/North N.S.W.	
Tibouchina urvilleana (DC.) Cogn./P.G.		Tibouchina urvilleana (DC.) Cogn./P.G.	
MYRTACEAE		MYRTACEAE	
Acmena smithii (Poir.) Merr. & Perry/Adelaide B.G.		Acmena smithii (Poir.) Merr. & Perry/Adelaide B.G.	
Eucalyptus microcarpa (Maiden) Maiden/P.G.		Eucalyptus microcarpa (Maiden) Maiden/P.G.	
ONAGRACEAE		ONAGRACEAE	
Circea cordata Royle/Missouri B.G.		Circea cordata Royle/Missouri B.G.	
Clarkia rubicunda (Lindl.) Lewis & Lewis/P.G.		Clarkia rubicunda (Lindl.) Lewis & Lewis/P.G.	
Clarkia unguiculata Lindl./P.G.		Clarkia unguiculata Lindl./P.G.	
Epilobium canum (Greene) Raven/Adelaide B.G.		Epilobium canum (Greene) Raven/Adelaide B.G.	
Epilobium ciliatum Raf./P.G.		Epilobium ciliatum Raf./P.G.	
Fuchsia hybrida Voss/P.G.		Fuchsia hybrida Voss/P.G.	
Gaura lindheimeri Engelm. & A. Gray./Adelaide B.G.		Gaura lindheimeri Engelm. & A. Gray./Adelaide B.G.	
Hauiya elegans DC./Missouri B.G.		Hauiya elegans DC./Missouri B.G.	
Lopezia semeiandra (Plitmann) Raven & Breedlove/ Missouri B.G.		Lopezia semeiandra (Plitmann) Raven & Breedlove/ Missouri B.G.	
Ludwigia peploides (Kunth) Raven/Milang S.A.		Ludwigia peploides (Kunth) Raven/Milang S.A.	
Oenothera stricta Ledeb./Murray Bridge S.A.		Oenothera stricta Ledeb./Murray Bridge S.A.	
PUNICACEAE		PUNICACEAE	
Punica granatum L./P.G.		Punica granatum L./P.G.	
TRAPACEAE		TRAPACEAE	
Trapa natans L. var. japonica Nakai/Japan		Trapa natans L. var. japonica Nakai/Japan	
GROUP 16		GROUP 16	
CELASTRACEAE		CELASTRACEAE	
Cassine papillosa (Hochst) Kuntze/Adelaide B.G.		Cassine papillosa (Hochst) Kuntze/Adelaide B.G.	
Euonymus japonica L./Adelaide B.G.		Euonymus japonica L./Adelaide B.G.	
OLACACEAE		OLACACEAE	
Ochanostachys amantacea Masters/Singapore B.G.		Ochanostachys amantacea Masters/Singapore B.G.	
Strombosia javanica Blume/Bogor B.G. Indonesia		Strombosia javanica Blume/Bogor B.G. Indonesia	
GROUP 17		GROUP 17	
ANACARDIACEAE		ANACARDIACEAE	
Mangifera indica L./Brisbane P.G.		Mangifera indica L./Brisbane P.G.	

TABLE 2. Continued.

<i>Pistacia atlantica</i> Desf./Adelaide B.G. CONNARACEAE CNN	M K V W P P V G L K K Y E T L S Y L S P I N / T S E E L L K E V Q Y L L L S K W I P
<i>Agelaea trinervis</i> (Llanos) Merr./Bogor B.G.Indonesia Connarus conchocarpus F.Muell./N.Qld MELIACEAE MEL	M K V W P P L G K K K F E T L S Y L L P L S T Q Q L A Q E V D Y L L L S G W V P M K V W P P L G K K K F E T L S Y L P D M T P K Q L A K E V N Y L L L N G W V P
<i>Cedrela sinensis</i> A.L.Juss./Adelaide B.G. RUTACEAE RUT	M Q V W P P V / I G K K K F E T L S Y L P P L S T E E L G K E V D Y L L R S G S V P
<i>Calodendrum capense</i> Thunb./Adelaide B.G. Correa pulchella Mackay ex Sweet/P.G. Flindersia collina Bailey/Adelaide B.G. SAPINDACEAE SAP	M K V W P T V G L K K K F E T L S Y L P P L S P E E L L K E V D Y L L L M G W V P M K V W P T V G M K K K F E T L S Y L P P L S P E Q L L K E V D Y L L L M G W V P M Q V W P P F G K K K F E T L S Y L P P L S P E E L L K Q V D Y L L K K G W I P
<i>Dodonaea viscosa</i> Jacq./Ouyen Vic. Harpullia pendula Planchon ex F.Muell./Adelaide B.G. SIMAROUBACEAE SMR	M K V W P T T G L K K K F E T L S Y L P T L S S E E L A K E V N Y L L L K G W I P M Q V W P P L G D K K F E T L S Y L P P L S L E Q L A K E I E Y L L L K G W I P
<i>Ailanthus altissima</i> (Miller) Swingle/Adelaide B.G. Kirkia wilmsii Engl./Adelaide B.G.	M Q V W P P T G K K K K F E T L S Y L P P L S P E E L L K E V D Y L L L K G S I P M K V W P T L G L K K K F E T L S Y L P P L S P E E L L K E I D Y L L L K G S I P
GROUP 18	
HALORAGACEAE HAL	
<i>Gonocarpus oreophyllus</i> A.E.Orchard/Canberra B.G. Haloragodendron bauerlenii (F.Muell.) A.E.Orchard/ Canberra B.G.	M Q V W P P L G L K K K F E T L S Y L P P L T T E Q L A K Q I D F L L L S K W V P M K V W P P V G L K K K F E T L S Y L P P L S P E E L A K Q V E F L L L Q G W V P
RHIZOPHORACEAE RHZ	
<i>Anisophyllea disticha</i> Baill./Kuching Forest Sarawak Bruguiera exaristata Ding Hou/Darwin N.T. Carallia brachiata (Lour.) Merr./Darwin N.T.	M M V W P P I G K K K Y F E T L S Y L P P L S S E A L A K E V E Y L I I K G W I P M L V W P P E G K K K F E T L S Y L P P L T T E Q L S K E V D Y L L L S K W V P M Q V W P P I G K K K F E T L S Y L P P L S S E E L A K E V D Y L L L S K W I P
GROUP 19	
GERANIACEAE GER	
<i>Erodium moschatum</i> L'Her./P.G. Pelargonium cucullatum (L.) Aiton/Adelaide B.G. MALPIGHIACEAE MLP	M K V W P P L G L K K K F E T L S Y L P P L T Q E E L G K E V D Y L L L S G W V P M E V W P P L G L K K K F E T L S Y L P P L T P E E L G K E V D Y L L L S G W V P
<i>Malpighia glabra</i> L./Adelaide B.G. Stigmaphyllon ciliatum (Lam.)Adr.Juss./Adelaide B.G. TROPAEOLACEAE TRP	M Q V W P T E G L S K F E T L S Y L P P L T D E E L I K E I D Y L L I S G W V P M M V W P T L G L S K F E T L S Y L P P L S A E E L L K E V E Y L L L S G W I P
<i>Tropaeolum majus</i> L./Adelaide B.G. ZYGOPHYLLACEAE ZYG	M Q V W P P L G L K K F E T L S Y L P P L T E V Q L A K E I D Y L I R S G W I P M K V W P T L G L K K F E T L S Y L P T L S E E G L K E I E Y L I I S G W I P
GROUP 20	
APOCYNACEAE APO	
<i>Mandevilla laxa</i> (Ruiz & Pavon) Woodson/Adelaide B.G. Vinca minor L./Adelaide B.G. ASCLEPIADACEAE ASC	M Q V W P P F I N K K K K Y E T L S Y L P P D L T P Q Q L A K E V D Y L L L K G W I P M K V W P P F L G L K K K F E T L S Y L P D L T E Q Q L A K E V D Y L L L K G W I P
<i>Hoya carnosa</i> R.Br./P.G. GENTIANACEAE GEN	M Q V W P P V G K K K K Y E T L S Y L P P L T N E Q L I K E I D Y L L L N K G I P
<i>Eustoma russellianum</i> G.Don/P.G. Orphium frutescens E.Meyer/Adelaide B.G. LOGANIACEAE LOG	M K V W P T E G L K K K F E T L S Y L P P L T T E Q L A K E I D Y L L L S G W I P M K V W P P V N M K K K F E T L S Y L P P L T T E Q L A K E V D Y L L L S G W V P
<i>Buddleia globosa</i> Hope/Mt Lofty B.G. S.A.	M K V W P P L G L K K F / Y E T L S Y L P P L T D E Q L L K E I D Y L L L S K W I P

TABLE 2. Continued.

Logania vaginalis (Labill.) F. Muell. / P.G.	M Q V W P T E G L K K F E T L S Y L P E L T E A Q L L K E V D Y L L L N G W I P
OLEACEAE	
Fraxinus excelsior L. / Adelaide B.G.	M K V W P P L G L K K Y E T L S Y L P D L S S E Q L L K E V D Y L L L S K W I P
Olea europaea L. / Adelaide B.G.	M K V W P P L G L K K F E T L S Y L P D L T S E Q L L K E V D Y L L L S K W I P
RUBIACEAE	
Coprosma repens A. Rich. / Adelaide B.G.	M K V W P P L G L K K Y E T L S Y L P P L T T E Q L L K E I E Y L I S M G W I P
Gardenia thunbergia L.f. / Adelaide B.G.	M Q V W P P T Y G K K K Y E T L S Y L P P L T T E Q L L K E I D Y L I S T G W V P
GROUP 21	
LAMIACEAE	
Phlomis cashmeriana Royle / Adelaide B.G.	M K V W P T E G L K K F E T L S Y L P P L T S E E L L K E V E Y L L L K N L I P
Teucrium flavum L. / Adelaide B.G.	M K V W P P E G G G V K K F E T L S Y L P P L S S E Q L I K Q V E F L L L K N L I P
VERBENACEAE	
Citharexylum myrianthum Cham. / Adelaide B.G.	M K V W P T E G L K K F E T L S Y L P D L S T E Q L L K E V D F L L L K K L I P
Vitex lucens Kirk / Adelaide B.G.	M K V W P T E G L K K F E T L S Y L P P L T T E / K E L L K Q V E Y L L L S K L I P
GROUP 22	
CONVOLVULACEAE	
Dichondra repens Forster & G. Forster / Adelaide B.G.	M Q V W P P V G L K K F E T L S Y L P P L S T E Q L L K Q V D Y L I R K G W V P
Ipomoea learii Paxton / P.G.	M Q V W P P I T / G L K K F E T L S Y L P D L S T E Q L L K Q V E Y L L N K G W V P
POLEMONIACEAE	
Cobaea scandens Cav. / P.G.	M Q / K V W P P L G L K K / E F E T L S Y L P P L T S Q / K E L A K E V D Y L L K K G W I P
Phlox paniculata L. / P.G.	M K V W P P L G L K K F E T L S Y L P P L T D E Q L A K Q V D Y L L L E G W V P
SOLANACEAE	
Anthocercis angustifolia F. Muell. / Mt Lofty S.A.	M Q V W P P I N M K K Y E T L S Y L P D L S E Q L L L E I E Y L L L S G W V P
Lycopersicon esculentum Millan / P.G.	M Q V W P P I N M K K Y E T L S Y L P D L T D E Q L L T E I E Y L L L K N G W V P
Nicotiana africana Merxm. / C.S.	M Q V W P P I N K K K K Y E T L S Y L P D L S Q E Q L L L E I E Y L L L K N S W V P
Nicotiana frugans Hook. / C.S.	M Q V W P P I N K K K K Y E T L S Y L P D L S Q E Q L L L E I E Y L L L K N G W V P
Nicotiana goodspeedii H. Wheeler / C.S.	M Q V W P P I N M K K K Y E T L S Y L P D L S Q E Q L L L E V E Y L L L K N G W V P
Nicotiana maritima H. Wheeler / C.S.	M Q V W P P I N M K K K Y E T L S Y L P D L S Q E Q L L L E V E Y L L L K N G W V P
Nicotiana petunioides (Grisebach) Millan / C.S.	M Q V W P P I N M K K K Y E T L S Y L P D L S Q E Q L L L E I E Y L L L K N G W V P
Nicotiana sylvestris Spegazzini & Comes / C.S.	M Q V W P P I N K K K K Y E T L S Y L P D L S Q E Q L L L E V E Y L L L K N G W V P
Nicotiana tabacum L. / C.S.	M Q V W P P I / Y N / G K K K Y E T L S Y L P D L S Q E Q L L L E V / I E Y L L K N G W V P
Petunia hybrida Vilm. / P.G.	M Q V W P P Y G K K K Y E T L S Y L P D L T D E Q L L K E I E Y L L L K N G W V P
Solanum aviculare G. Forster / P.G.	M Q V W P P I N K K K Y E T L S Y L P D L T D E Q L L T E I E Y L L L K N G W V P
Solanum chippendalei Symon / Alice Springs N.T.	M Q V W P P I G K K K Y E T F S Y L P D F S D E Q L L A E I D Y L L L S G W V P
Solanum cinereum R.Br. / Wilpena S.A.	M Q V W P P I N / G K K K Y E T F S Y L P D F S D E Q L L A E I D Y L L L S G W V P
Solanum macrocarpon L. / P.G.	M Q V W P P I N K K K K Y E T L S Y L P D L S D E Q L L L E I D Y L L L R S G W V P
Solanum marginatum L.f. / P.G.	M Q V W P P I N K K K K Y E T L S Y L P D L S D E Q L L L E I D Y L L L R S G W V P
Solanum melongena L. / P.G.	M Q V W P P I N K K K K Y E T L S Y L P D L S D E Q L L L E I D Y L L L R S G W V P
Solanum nigrum L. / P.G.	M Q V W P P I N K K K K Y E T L S Y L P D L S D E Q L L L E I D Y L L L R S G W V P
Solanum oligacanthum F. Muell. / P.G.	M Q V W P P I E / N K K K Y E T F S Y L P D L S D E Q L L L A Q I D Y L L R S G W V P
Solanum petrophilum F. Muell. / Alice Springs N.T.	M Q V W P P I G K K K K Y E T F S Y L P D F S D E Q L L A E I D Y L L L S G W V P
Solanum quadriloculatum F. Muell. / Alice Springs N.T.	M Q V W P P I N K K K K Y E T F S Y L P D F S D E Q L L A E I D Y L L R S G W I / V P
Solanum trilobatum L. / P.G.	M Q V W P P I N K K K K Y E T L S Y L P D L S D E Q L L L E I D Y L L R S G W V P
Solanum tuberosum L. / P.G.	M Q V W P P I N K K K K Y E T L S Y L P D L S D E Q L L L E I D Y L L R S G W V P
Solanum violaceum R.Br. / P.G.	M Q V W P P I N K K K K Y E T L S Y L P D L S D E Q L L L E I D Y L L R S G W V P
Solanum viridifolium Dunal / P.G.	M Q V W P P V G K K K Y E T F S Y L P T L S D E Q L L L E I D Y L I R S G W I P
GROUP 23	

TABLE 2. Continued.

CHRYSOBALANACEAE	CHB	Chrysobalanus icaco L./Lyon Arboretum Hawaii	M K V W P T T G L L K K F E T L S Y L P P L T T P Q L A K E V D Y L L L S G W V P P
		Parinari nonda F.Muell./CSIRO Darwin N.T.	M K V W P T V G K K K F E T L S Y L P P L T T P Q L A K E V D Y L L L T G W I P
CORIARIACEAE	CRR		
Coriaria nepalensis Wallich/Melbourne B.G.			M Q V W P P I G L K K Y E T L S Y L P P L S S A E L A K E V D Y L L S K G W V P
CROSSOSOMATACEAE	CRO		
Crossosoma californica Nutt./Rancho Santa Ana B.G.			M K V W P P L G L K K F E T L S Y L P P L T P E E L G K E I D Y L L L K G W V P
ELAEAGNACEAE	ELE		
Elaeagnus pungens Thunb./Adelaide B.G.			M Q V W P P Y G K K K K F E T L S Y L P P L T T E Q L A K E V D Y L L R N G W V P
Shepherdia argentea Nutt./P.G.			M Q V W P P Y G K K K K F E T L S Y L P P L T T E S L A K E V D Y L L L N G W V P
EUPHORBIACEAE	EUP		
Acalypha wilkesiana Muell. Arg./Adelaide B.G.			M Q V W P P V G S K Y F E T L S Y L P P L T D E S L A S Q V E Y L L L N G W V P
Glochidion ferdinandii Muell. Arg./Adelaide B.G.			M K V W P T F G L / S K K F E T L S Y L P P L T S D D Q L A K E V E Y L L L S G W V P
Ricinus communis L./P.G.			M Q V W P P V / L G S K K F E T L S Y L P P L T T E S L A S E V E Y L L L N G W V P
GOODENIACEAE	GOD		
Goodenia ovata Smith/Mt Lofty B.G. S.A.			M K V W P P L G L K K K Y E T L S Y L P P L D S E E / G L A K E V D Y L L L N K K W V P
Scaevola albida (Smith) Druce/Belair S.A.			M K V W P P T G L K K K Y E T L S Y L P P L S T D E L A K E I D Y L L L N K K W V P
HYDROPHYLLACEAE	HYD		
Nemophila menziesii Hook. & Arn./C.S.			M K V W P P E N N K K M Y E T L S Y L P P D L T T E Q L A K Q I D Y L L I N K K W I P
Phacelia tanacetifolia Benth./C.S.			M K V W P P I N L K K Y E T L S Y L P P L T T E Q L A K E V D Y L L I N K K W V P
LECYTHIDACEAE	LCY		
Couroupita guianensis Aublet/Lyon Arboretum Hawaii			M K V W P P S N N K K K F E T L S Y L P N L T D E E L L K E V E Y E L L L G W V P
Lecythis pohlili Berg/Lyon Arboretum Hawaii			M K V W P P S N N K K K F E T L S Y L P N L S D E Q L L K E V E Y I L L S G W V P
LOASACEAE	LOA		
Mentzelia lindleyi Torrey & A.Gray/C.S.			M Q V W P P I G L K K K Y E T L S Y L P P L T T D Q L A K E I E Y L I I S G W V P
NELUMBONACEAE	NEL		
Nelumbo nucifera Gaertner/Adelaide B.G.			M Q V W P P T G K K K F E T L S Y L P P L S P E E L A K E V D Y L L L M G W I P
PIPERACEAE	PIP		
Peperomia polybotrya Kunth/P.G.			M Q V W P P V G K K K K F E T L S Y L P P L T Q E G L L K E I E Y L L I S G W I P
Piper nigrum L./Adelaide B.G.			M Q V W P P E G K K K F E T L S Y L P T / P L S Q E G L A K E I E Y L L M S G W I P
PLUMBAGINACEAE	PLB		
Limonium gmelinii Kuntze/Adelaide B.G.			M M V W P P E G L K K K F E T L S Y L P P L S S E A L A N E I A Y L I S M T W V P
Plumbago auriculata Lam./Adelaide B.G.			M M V W P P E G K V K F E T L S Y L P P L S S E A L A N E I E Y L L S M T W I P
POLYGONACEAE	PLG		
Fagopyrum esculentum Moench/C.S.			M Q V W P P S G K K K K F E T L S Y L P P L S E K D L I A Q V N Y L L V K G W V P
Rheum raphaniticum L./P.G.			M Q V W P P E G K L K F E T L S Y L P P L S E A E L L A Q I N Y L L I N G W V P
Rumex obtusifolius L./P.G.			M Q V W P P E G K I K F E T L S Y L P P L S E A D L L A Q I N Y L L I N G W V P

TABLE 2. Continued.

PROTEACEAE		PRT
Adenanthos sericeus	Labill./Canberra B.G.	M Q V W P P I N K R K F E T L S Y L P L S P E A L A K E V D Y L L R K G W V P
Brabejum stellatifolium	L./Kirstenbosch B.G.	M K V W P L P G L L / G K K K F E T L S Y L P L S P E Q L A K E V N Y L L L N G W V P
Darlingia darlingiana	(F.Muell.) L. Johnson/N.Qld	M Q V W P P L G K K K F E T L S Y L P L S P E A L A K E V D Y L L L K G W V P
Embothrium coccineum	Forst.et f./Wakehurst B.G. U.K.	M K V W P P L G L K K F E T L S Y L P L S P E G L A K E V D Y L L I R K G W V P
Isopogon latifolius	R.Br./P.G.	M Q V W P P V G K K K F E T L S Y L P L S P E Q L K K E V D Y L L L R K G W V P
Knightia excelsa	R.Br./Otari B.G. N.Z.	M K V W P P L G L K K F E T L S Y L P L S P E E L A K E V D Y L L L L K G W V P
Leucadendron laureolum	(Lam.) Fourc./Adelaide B.G.	M Q V W P P V G M L K F E T L S Y L P L S S E Q L A K E I D Y L L L R K G W V P
Lomatia arborescens	L.Fraser & J.W.Vickery/ Canberra B.G.	M Q V W P P V G K K K F E T L S Y L P L S S E I L A K E V D Y L L L R K K W V P
Lomatia dentata	R.Br./Wakehurst B.G. U.K.	M K V W P P L G D K K K F E T L S Y L P L S S E I L A K E V D Y L L L R K G W V P
Macadamia ternifolia	F.Muell./Adelaide B.G.	M Q V W P P L G K K K F E T L S Y L P L S P E Q L A K E I N Y L L L I K G W V P
Persoonia chamaepeuce	Lhotsky ex Meissner/ Canberra B.G.	M K V W P P K G I P K F E T L S Y L P L S S E E L A K E V D Y L L L L K G W V P
Protea repens	(L.) L./P.G.	M Q V W P P V G L K K K F E T L S Y L Q P L S S I E L A K E V D Y L L I K K W V P
Telopea mongaensis	Cheel/Canberra B.G.	M K V W P P L G L K K F E T L S Y L P P L T P E I L A K E V D Y L L R K K W V P
Toronia toru	(A.Cunn.) L. Johnson & B.Briggs/ Otari B.G. N.Z.	M Q V W P P K G I P K F E T L S Y L P P L S S E E L A K E V D Y L L L K G W V P
RHAMNACEAE		
Rhamnus cathartica		RHM
Ceanothus thyrsiflorus	Eschsch./Adelaide B.G.	M Q V W P P L G L K K F E T L S Y L P D L T S D Q L A K E I D Y L L I K G W I P
Rhamnus cathartica	L./P.G.	M Q V W P P L G I K K F E T L S Y L P P L T N Q Q L I K E V E Y L I I K G W I P
THYMELAEACEAE		
Daphne odora		THY
Pimelea physodes	Hook./P.G.	M Q V W P P V G L K K F E T L S Y L P P L T S E Q L A K E I D Y L L I R N K W I P
		M Q V W P P T N N K K F E T L S Y L P P L S T V Q L A K E I E Y L I R N K W I P
VITACEAE		
Parthenocissus quinquefolia		VIT
Vitis vinifera	L./Adelaide B.G.	M K V W P T E G L I K K F E T L S Y L P P L S S E E L G K E V D Y L L L K K W V P
		M K V W P T E G L M K K F E T L S Y L T P L S Y E Q L A / R K E V D Y L L L K K W I P
M O N O C O T Y L E D O N S		
Sagittaria graminea		ALI
APONOGETONACEAE	APG	M K V W P V E G L K K F E T L S Y L P P L T T E Q L L K Q I E Y L L R N G W I P
Aponogeton distachyos	L.f./Mylor S.A.	M Q V W A P E G L K K F E T L S Y L P P L T S E Q L L K E V D Y L I I K G W I P
ARACEAE	ARA	M Q V W P P E G L K K F E T L S Y L P P L S D E A L A K E V D Y L I I K G W I P
Arum maculatum	L./P.G.	M Q V W P P L G L K K F E T L S Y L P P L S N I N L S K E V D Y L L L N G W I P
Philodendron giganteum	Schott/P.G.	M K V W P I / V E G L A / K K F E T L S Y L P P L S I E A L L K Q I E Y L L L S G W V P
ARECACEAE	ARE	M K V W P I E G H / K K F E T L S Y L P P L P I E A L L K Q I E Y L L L S G W I P
Brahea armata	S.Watson/Adelaide B.G.	
Chrysalidocarpus lutescens	(Bory)Wendl./ Adelaide B.G.	
COMELINACEAE		
Commelina coelestis		CMM
Rhoeo discolor	Hance/P.G.	M K V W P I E G V K K F E T L S Y L P P L T P E I L L K Q V D Y L L L S G W V P
LILIAEAE	s.l. LIL	M K V W P I E G K K K F E T L S Y L P P L S N T I L L K Q I D Y L I L S K W V P
Asparagus falcatus	L./Adelaide B.G.	M K V W P I A G K K K F E T L S Y L P P L S I E Q L I K Q I E Y L I L S S W I P

TABLE 2. Continued.

Aspidistra elatior Blume/Adelaide B.G.	M Q V W P I E G V K K F E T L S Y L P P L S I E Q L I K Q I D Y L L R N G W I P
Liriope spicata Lour./Adelaide B.G.	M K V W P I E G M K K F E T L S Y L P D M T T V Q L I K K I E Y L L N K W I P
Ruscus hypophyllum L./Adelaide B.G.	M K V W P I E G V K K F E T L S Y L P D M T I E Q L I K Q I E Y L I K N K W I P
ORCHIDACEAE ORC	
Cymbidium X floribundum Lindley X C.lowianum (Reichb.f.)/see ref. G.C.Martin	M Q V W P I E G V K K F E T L S Y L P T/PL ST/PE A L L K Q I D Y L L R S G W I P
POACEAE POA	
Hordeum vulgare L./C.S.	M Q V W P I E G I K K F E T L S Y L P P L S T E A L L K Q V D Y L I V K K W V P
Lolium perenne L./C.S.	M Q V W P I E G I K K F E T L S Y L P P L S A E A L L K Q I D F L I R K K W V P
Phragmites australis (Cav.)Trin. ex Steudel/ River Torrens S.A.	M Q V W P I E G I K K F E T L S Y L P P L T T E D L L K Q I E Y L I V K K W I P
Piptatherum miliaceum (L.)Cosson/P.G.	M Q V W P I E G I K K F E T L S Y L P P L S V E A L/FL K Q I E F L I V K K W V P
Triticum vulgare L./C.S.	
SMILACACEAE SML	M Q V W P I E G I K K F E T L S Y L P P L S T E A L L K E V D Y L I V K K W V P
Eustrephus latifolius R.Br./Canberra B.G.	M K V W P/IP/IEG M K K F E T L S Y L P D L S V E Q L I K Q I E Y I L I N K W V P
Geitonoplesium cymosum (R.Br.)A.Cunn. ex Hook./ Canberra B.G.	M K V W P I E G L K K F E T L S Y L P P L T V D Q L L K Q I D Y I L I S K W I P
G Y M N O S P E R M S	
EPHEDRACEAE EPH	
Ephedra fragilis Desf./Adelaide B.G.	M L V W P N K D N I K F E T L S Y L P T L S Q E A L I K Q I D Y L L S K W V P
GINKGOACEAE GNK	
Ginkgo biloba L./P.G.	M Q V W N P Y N N K K F E T L S Y L P P L S P E G L L K E I D Y L L N N G W V P
CYCADACEAE CCD	
Macrozamia communis L.Johnson/Bateman's Bay N.S.W.	M Q V W P P Y N N K K F E T L S Y L P P L S K P Q L L E I D Y L L G N K W I P
TAXODIACEAE TXO	
Metasequoia glyptostroboides Hu&Cheng/Adelaide B.G.	M Q V W P P Y N N K K W E T L S Y L P P L S P E Q I L K Q V I Y I L K N K W V P
WELWITSCHIAEAE WLW	
Welwitschia mirabilis Hook.f./C.S.	MQ/LV W P N KG/NL L K F E T L S Y L P D L T P E Q L L K Q I D Y L L N K W V P

angiosperm ancestors. All three members of Acanthaceae that were tried failed with symptoms like these, as did four out of six species from Caesalpiniaceae.

Finally, an entirely different sort of failure occurred with four species, all hemiparasites from the putatively related families Santalaceae and Loranthaceae. These species had abnormally high amounts of phenolics, but it seems unlikely that failure can be attributed to them or to any of the other causes mentioned above. The preparations always yielded abnormally high amounts of plastocyanin but no trace of rubisco-SSU. Plastocyanin, a chloroplast protein, has a molecular weight sufficiently close to rubisco-SSU that it occurs, occasionally, as a small contaminant detected during sequencing. It could be identified by its sequence but, except in these two families, it was so weak that it disappeared after about seven positions. The strength of the plastocyanin sequence in all four of these hemiparasites suggests that the absence of rubisco-SSU could not be ascribed to some general difficulty like proteases, but might reflect an unusual, perhaps facultative, photosynthetic system.

GENERAL REMARKS ABOUT THE SEQUENCES OVERALL VARIATION AND INVARIANT SITES

A summary of the variation that we have observed is given in Table 3; the amino acids most commonly observed are in the top line.

The rubisco-SSU gene includes two introns, the first of which is inserted before the codon that determines amino acid 3. It determines valine and this, like tryptophan at position 4, is invariant, the two codons carrying the signal to cut the end of the intron (Berry-Lowe et al., 1982). These invariant residues were useful early signals that the correct protein fraction had been chosen. Within the first 40 amino acids, proline always occurs at position 5 and/or 6, at position 19 and/or 20, and at position 40. These three regions correspond to bends in the tertiary structure of the molecule. Chapman et al. (1988) have indicated that between the first and second bend there is alpha-helix and thereafter beta-sheet. There is an almost invariant region from amino acids 13 to 18, a region that makes contact with one of the large subunits (Chapman et al., 1988; Knight et al., 1989). The only variation we have found in this region is the substitution of phenylalanine for leucine at position 15 in five species of *Solanum* (Martin et al., 1986). These same species also have phenylalanine substituted for leucine at the almost invariant position

21. The simultaneous occurrence of two very rare substitutions indicates a causal connection. Hydrophobic bonding between the two positions may stabilize the bend at position 19 and, because these species are inhabitants of very hot and arid regions, this may have been a factor in natural selection.

HETEROGENEITY WITHIN SPECIES

The first reports of rubisco-SSU sequences (Strobaek et al., 1976) were for the N-terminal amino acids in species of *Nicotiana*, and these showed heterogeneity at positions 7 and 8 in tobacco. We have also found it at position 30. These heterogeneities are undoubtedly associated with the amphidiploid origin of tobacco and led to the expectation that heterogeneity would be fairly common, not only because about one-third of plant species are polyploid, but also because in diploids gene duplications are frequent. We may not have detected some heterogeneities (for example, those involving serine, which gives a weak signal), but we did detect 34 species with one heterogeneity, 11 with two, and 4 with three. The demonstration by Pichersky et al. (1986) that in tomato there were at least three different DNA messages for rubisco-SSU, all with the same N-terminal amino acid sequence, suggests that selection acts strongly to preserve primary amino acid structure. There are at least eight different genes encoding rubisco-SSU in petunia (Lamb & Fitzmaurice, 1986); for this reason, when we prepared protein from that species, we used a mixture of equal quantities of leaves from four morphologically different varieties, with the aim of finding heterogeneities (Martin & Dowd, 1984b). The sequence was of high quality but no heterogeneity was detected. Likewise, we chose to study *Rhoeo discolor* because it is a complex interchange heterozygote for all chromosomes and might therefore be heterozygous for rubisco-SSU, but we detected no heterogeneity. Heterogeneities that were found presented no problem for the computer analysis.

INSERTIONS

Only two examples of additional amino acids in the N-terminal sequence have been found. Both species of Epacridaceae that we studied had an additional isoleucine between normal positions 9 and 10. *Teucrium flavum* (Lamiaceae) had two additional glycines, probably between the same two positions. These insertions, while clearly of taxonomic significance, have been ignored during data processing.

THE N-TERMINUS

Haslett et al. (1976) reported that the N-terminus of rubisco-SSU was "frayed," some molecules seeming to have methionine in position 1 while others are without it. This is the situation that we have encountered in the vast majority of species, the effect being that at every position two amino acids are recorded, the correct one and the next one. Usually the two signals are approximately equal, especially when the protein is of highest quality. This property is helpful in that it provides a second opportunity for identification and is useful for identifying minor contaminating proteins whose residues appear only once, but probably means that it is more difficult to obtain long sequences because attenuation of the signal occurs earlier.

All those species for which nucleic acid sequences have been reported were also studied by us and all show fraying. Because the nucleic acid sequences show the N-terminus to be methionine, there is no doubt. The signals we obtained were not typical for either methionine or its sulfone derivative. Whether the derived amino acid is obtained by dansylation (in manual sequencing) and identified by TLC, or is the PTH derivative from automatic sequencing and identified by HPLC or TLC, the N-terminal amino acid moves differently from methionine; therefore, we conclude that it is a modified form of that amino acid.

Two exceptions to the above generalization have been encountered. In 10 out of 11 species of the Onagraceae the N-terminus is phenylalanine, the only variations from methionine known, and in these there is no sign of fraying. In six other species the N-terminal amino acid is methionine (and gives the normal signal for PTH-methionine), but there is no sign of fraying, the difference from the majority of species being sharp and unmistakable; these are two members of Papaveraceae (*Papaver orientalis* and *Eschscholtzia californica*), two from Pedaliaceae (*Sesamum indicum* and *Ceratotheca triloba*), *Vitex lucens* (Verbenaceae), and *Mentzelia lindleyi* (Loasaceae). Any hypothesis to explain fraying must account for these exceptions, and we believe that they exclude artifacts arising from techniques of protein production or sequencing. Any hypothesis must also account for the modification of methionine and the equality of the two forms of the protein. We therefore dismiss as unlikely hypotheses relying on inefficient shortening of the protein either as it passes through the chloroplast membrane or after entry.

It is known that rubisco-SSU forms dimers (Roy et al., 1978), and we suggest that this may be

through formation of disulphide bridges between the N-terminal methionines of two SSU molecules. This might occur in vivo if the enzyme model of Chapman et al. (1988) is correct, but is more likely an in vitro event if the different model of Knight et al. (1989) is correct. If dimers are formed, S-carboxymethylation at pH 8.6 would not break an inter-methionine bond, but we suggest that the dimer does fall apart so that dimethionine is on one chain and no methionine on the other. This hypothesis would account for all phenomena except for non-fraying species, which presumably do not naturally form dimers.

METHODS OF DATA ANALYSIS

Before computer analysis, amino acid sequences were converted to inferred nucleotide sequences using the genetic code. Usually this could be carried out after inspection of, for example, all the sequences in a Group so that the most parsimonious choices of codons could be made. A standard was chosen at sites where substitution was silent. Although a program was available (Martin et al., 1983), usually the path was obvious and computing unnecessary. Thus the unit of length in phylogenetic trees is an inferred nucleotide difference (i.n.d.).

The number of dichotomizing trees (phylogenetic Steiner trees) connecting N taxa is $1 \times 3 \times 5 \dots (2N-5)$. The principle of analysis is that the length of every possible tree is calculated and the shortest tree is chosen as the most probable. This agrees with the parsimonious hypothesis that evolution has proceeded by the shortest route. However, because the total number of possible trees increases very rapidly, i.e., when increasing from $N-1$ to N taxa it increases $(2N-5)$ -fold, it is not always possible to consider every tree.

Except during the final stages of this project, the program that we used was MINTREE, the "branch and bound" program of Hendy and Penny (1982). With a Vax 785 computer the usual limit for simultaneous analysis was 12 taxa. This limit could be extended to about 15 with a supercomputer (Cyber 205), but the trouble and expense precluded useful work. Although MINTREE has now been superseded, its co-program ANALYZE is still used because it possesses efficient routines for obtaining ancestral sequences and internodal lengths.

Most of the analyses have been carried out using HENNIG86 (Farris, 1988) and a personal computer (Microbyte 230). This system is about two

orders of magnitude faster than the above and has the additional advantage that an analysis can be left running for days or weeks. The principles of its algorithms have not been published, but the time of its release suggests a possible connection with a published letter by Johnson (1987). HENNIG86 offers a number of options, which comparison with MINTREE using the same sets of data suggest are reliable; in order of preference we have used implicit enumeration (*ie**); *ie* followed by *bb* (branch swapping); *mhennig* followed by *bb*. Beyond the number of taxa that can be handled by MINTREE or the *ie* option of HENNIG86, correct solutions cannot be guaranteed.

A further advantage of HENNIG86 is that it includes a program for successive weighting, which often reduces the result to one or very few trees. This usually eliminates the need to derive consensus trees, a process that we have found unsatisfactory (Martin & Dowd, 1989). Finally, HENNIG86 derives the *ci* (consistency index) (Carpenter, 1988) and *ri* (retention index) which we record with each figure of a tree. In a personal communication, Farris defines these as follows: if *r* and *m* denote, respectively, the smallest and greatest number of steps that a character can require on any tree, and *s* denotes the number of steps that character requires on a considered tree, then *c.i.* is r/s and *r.i.* is $(m-s)/(m-r)$.

MINTREE uses data such that each of the four nucleotides is entered as 1, 2, 4, or 8, which allows the counting of ordinary differences and also of heterogeneities; no matter what variation occurs at a site, it can be recorded as a sum that is always different for different combinations. Provided there are no heterogeneities, HENNIG86 can use the same notation (using the nonadditive option); if there are heterogeneities, they must be recorded by inserting additional taxa. This is satisfactory if there is only one variable site within a taxon when only two taxa need to be recorded. Assumptions of linkage must be made, however, if there is more than one variable site but only two taxa are to be recorded. This problem becomes increasingly important as an analysis progresses from using raw data to derived ancestral sequences for families and then Groups because, in these, heterogeneities may be numerous. We have therefore used alternative strategies. The first is inserting additional taxa as just described and accepting the result if the different versions of a taxon cluster without interruption.

The second strategy is to use binary coding for nucleotides; e.g., 1000 for A, 0100 for G, 1100 for A and G, 0010 for C, and so forth. In con-

junction with HENNIG86, MINTREE notation is slower than binary notation, which is therefore advantageous. As would be expected, binary gives a tree length double that using MINTREE notation but it is seldom exact. If inexact, the length is always less than double, and there is a loose relationship between the deficit and the number of heterogeneities. Because the details of HENNIG86 have not been published, we have been cautious about choosing between these alternatives and have done all Group analyses with HENNIG86 using both notations.

We structured our investigation such that the majority of families were represented by at least two species from different genera. If we accept that taxonomy is seldom wrong when placing genera within families (Heywood, 1978), then we have an empirical way of judging the merits of the two notations. Omitting families that have either a single representative or are multiply represented, all Groups have been analyzed using both notations and, from the minimal trees recorded, we have chosen the best as judged by pairing of representatives of families. In eight Groups both methods had the same best tree, in six binary gave the best, and in eleven MINTREE notation gave the best. In the section "Analyses Within Groups" we have therefore used the taxonomically best minimal tree no matter which notation was used to derive it. However, in later sections we have used binary notation exclusively because it is quicker and more convenient.

Among best trees, 79% of families showed correct pairing of its members. When judging this result, it should be remembered that a single misplaced species will often result in the failure of pairing of representatives of two families. While a few such occurrences may be the result of incorrect taxonomy, the remainder are presumably caused by convergent evolution. The details can be seen in the figures for the Groups.

ANALYSES WITHIN GROUPS EXPLANATION OF THE FIGURES

The figures are drawn to scale, which is indicated by the length of one inferred nucleotide difference (i.n.d.). Only lengths have meaning, not angles. Usually at least two trees have been derived for each Group, one using sequences of individual species and the second using derived familial nodes. If the first shows congruent grouping of putative members of families, then the second is not needed.

Disruption of familial grouping is often caused by the sole representative of another family or by a member of a multiply represented family. The second tree is derived from familial nodes and singletons, is drawn with a different scale, and uses only the three-letter familial abbreviation of Weber (1982), which is given in Table 2. When appropriate, trees of multiply represented families or genera are also given. In later analyses, Group nodes, and often one or two others, will be used; each of these is numbered.

The Base of the Angiosperm Tree

Before detailed analysis of Groups began, many analyses were done using the five gymnosperms, representatives of the monocotyledons and of Groups 1, 2, and 3 which were most likely, on taxonomic grounds, to be near the root of the dicotyledon tree. The angiosperm family closest to the gymnosperms was Schisandraceae. Figure 1 shows the junction of the gymnosperms, Schisandraceae, and the other angiosperms. The derived sequence of this node has been used as “Base” in all subsequent Group analyses.

It will be noted that Figure 1 is different from, and taxonomically more satisfactory than, the equivalent figure of Martin and Dowd (1989). Since then a sequence of *Welwitschia* has been obtained and this paired with *Ephedra* between the angiosperms and the other gymnosperms. Three attempts to study *Gnetum* were made but all failed with symptoms suggesting strong leaf protease activity.

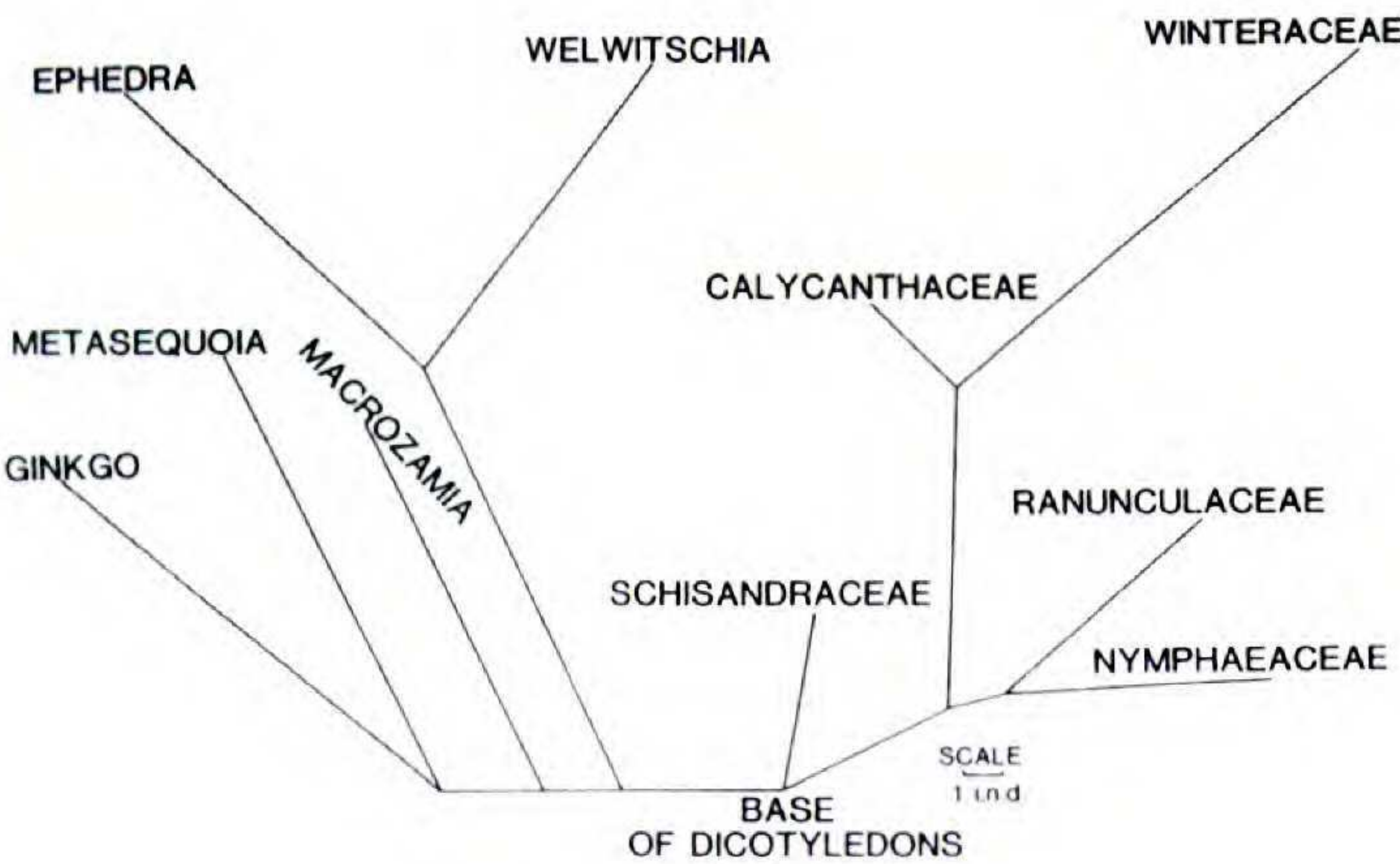


FIGURE 1. Five gymnosperms analyzed with familial nodes of five angiosperm families from Groups 1, 2, and 3. The ancestral sequence derived for the junction of Schisandraceae has been used as an outgroup for analyzing the Groups of dicotyledons.

Group 1. An attempt to study *Hedycarya* (Monimiaceae) having failed, *Peumus* was left as a singleton, which was therefore omitted from Figure 2a but included in Figure 2b. Correct pairing and grouping occurs in all families except Aristolochiaceae for which a derived familial node is shown in Figure 2b. As indicated above, Schisandraceae is nearest to Base with a rather long gap to the remainder.

Group 2. In contrast to the straightforwardness of the previous Group, Group 2 has presented problems that arise from the great intrafamilial variation of Ranunculaceae and Papaveraceae, the trees for which are shown in Figure 3b and c. The two derived familial nodes are shown with the rest of the Group in Figure 3a. We interpret Ranunculaceae and Papaveraceae to be ancient angiosperm families, and only with some misgivings have we adhered to our acceptance of current taxonomy at the levels of family and below. This is especially so for Papaveraceae, but splitting off Fumariaceae creates more problems than it solves.

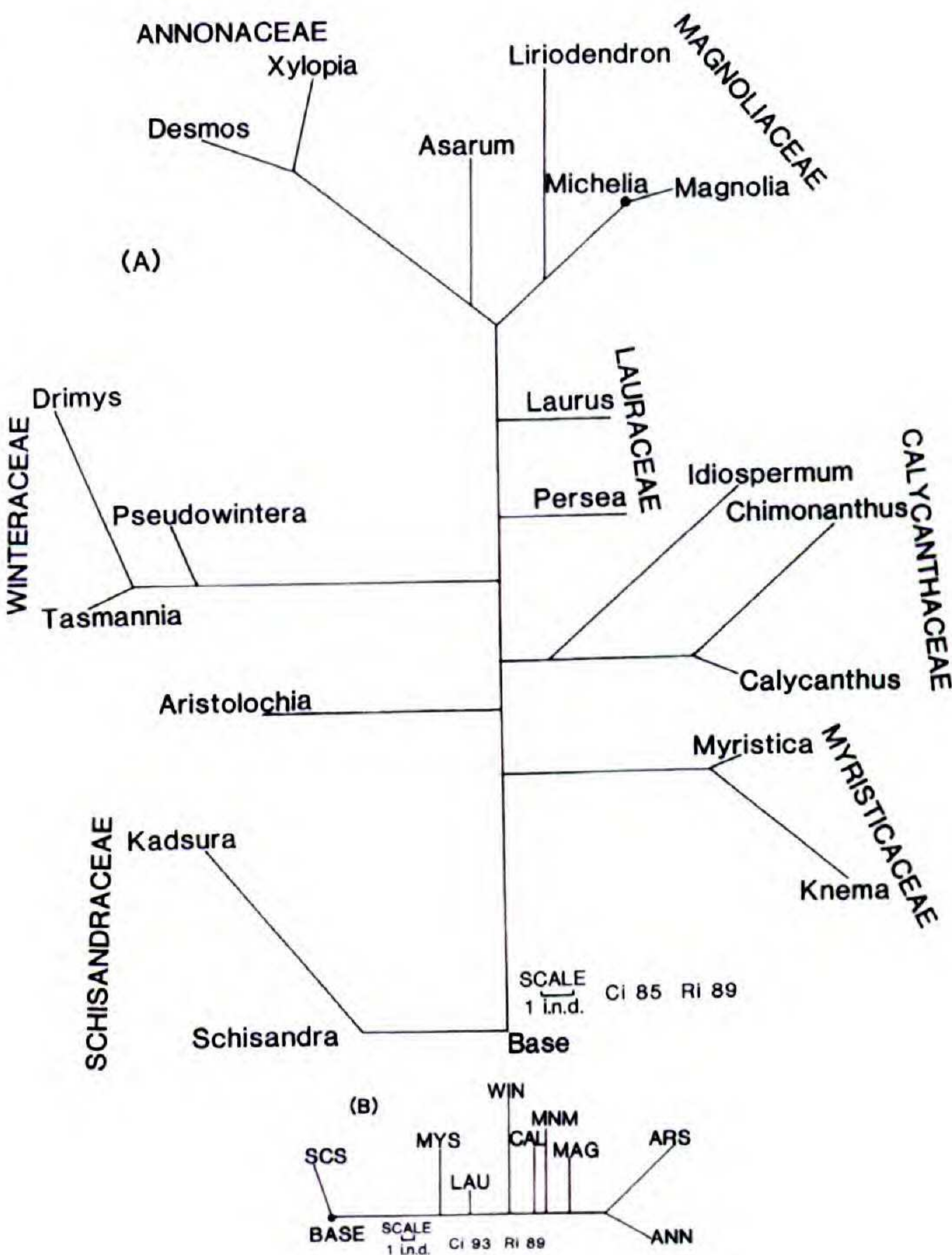


FIGURE 2.—(A). Group 1, omitting the single representative of Monimiaceae.—(B). Tree of family nodes for Group 1.

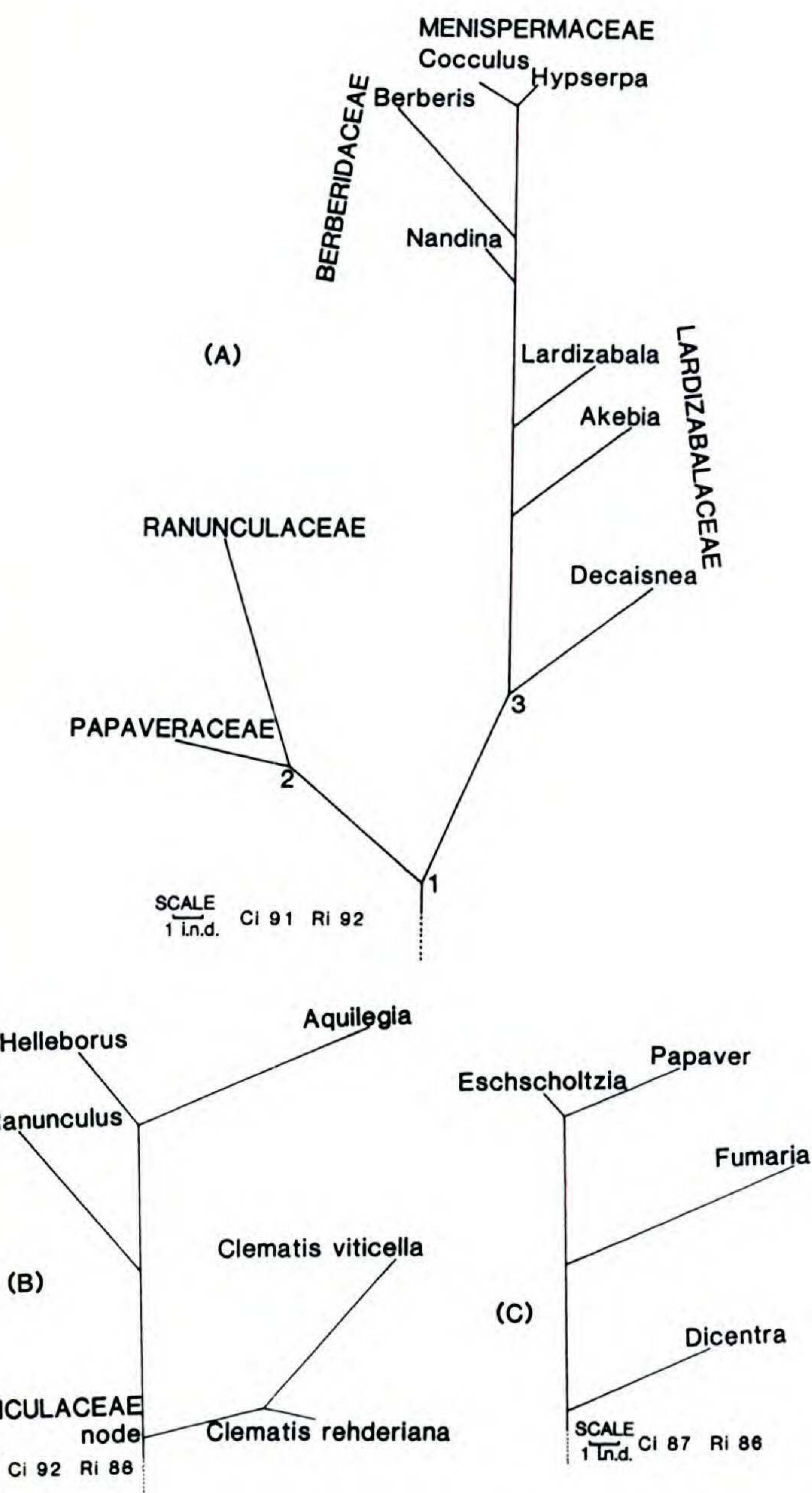


FIGURE 3. (A) Group 2, with family nodes, derived from (B) and (C) for Ranunculaceae and Papaveraceae s.l.

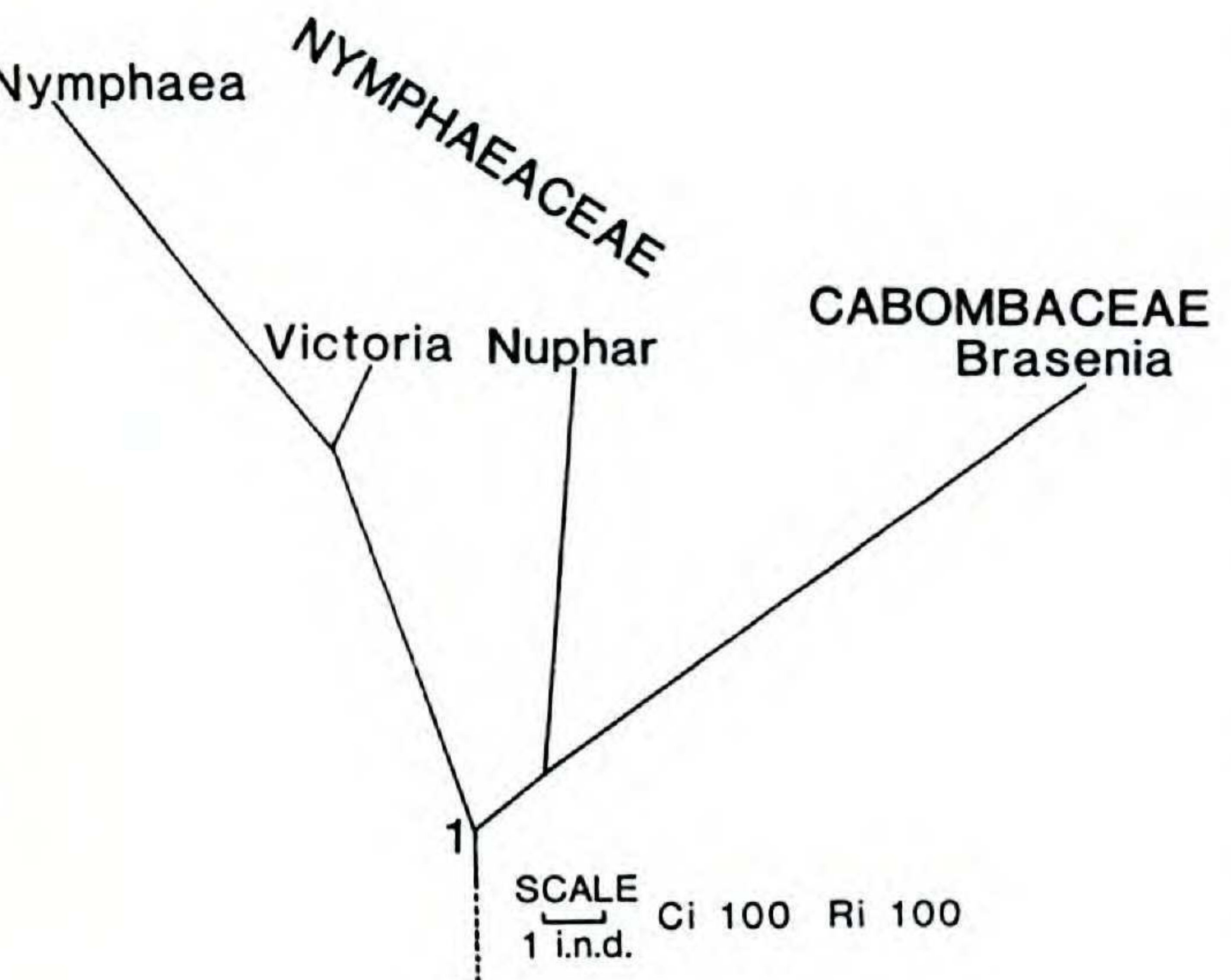


FIGURE 4. Group 3.

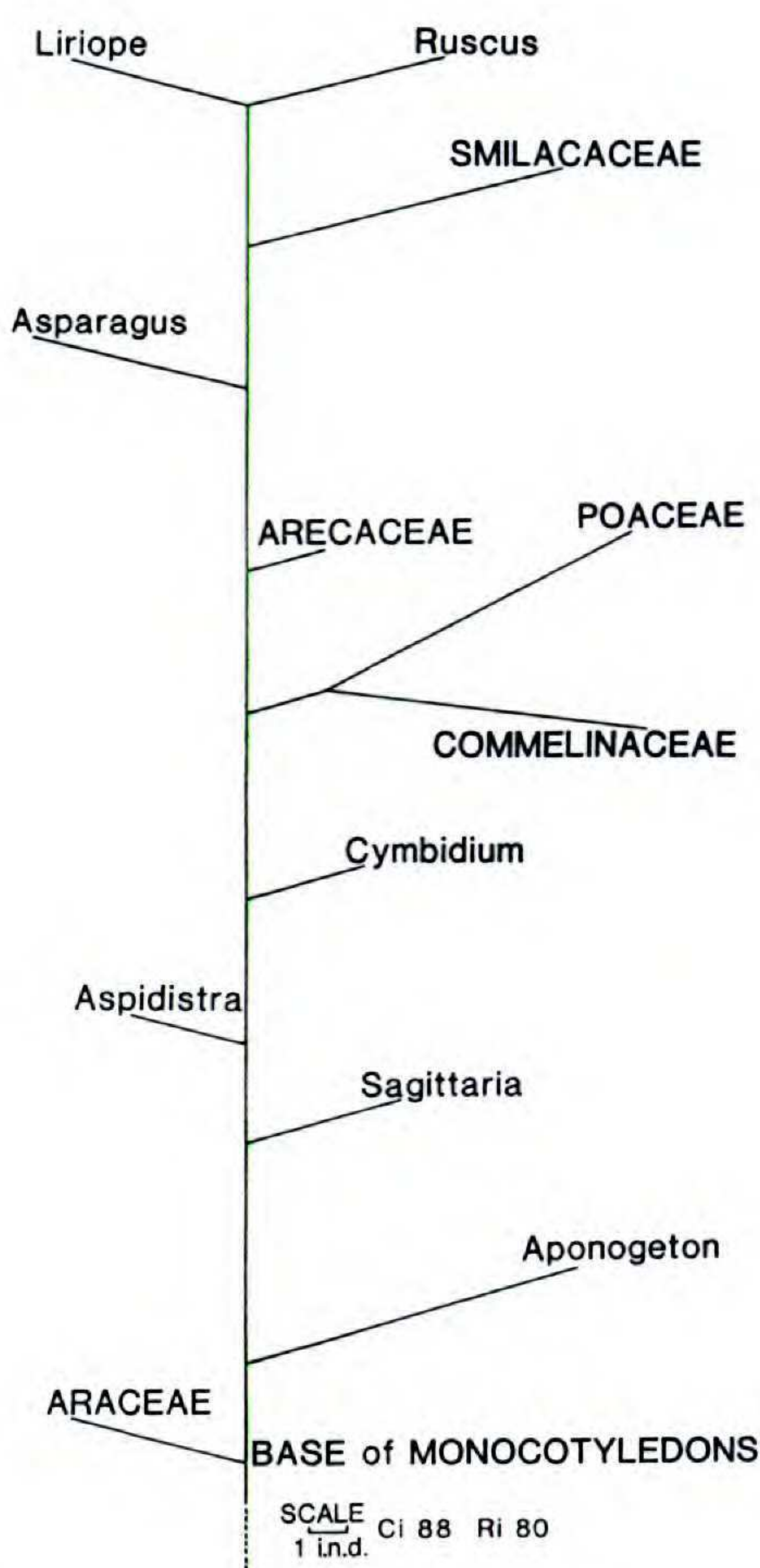


FIGURE 5. The monocotyledons that have been studied with some families represented by their nodes.

Group 3. Failure with a species of *Cabomba* has left *Brasenia* as a singleton that does not separate from the three members of *Nymphaeaceae*; however, the internode is so short that we draw no significance from it (Fig. 4).

Piperaceae, Nelumbonaceae, and the monocotyledons. Piperaceae and Nelumbonaceae have not been placed in a Group because in both cases opinions differ among the four phylogenies considered when we nominated Groups. All place them in either our Group 3 or Group 1, so it is appropriate to carry out a joint analysis with the members of these two Groups, and at the same time to consider the links with the monocotyledons. Although no species have been added to those reported earlier (Martin & Dowd, 1989), the sequences have been reanalyzed with HENNIG86. To reduce the number of taxa to be compatible with the *ie* option, familial nodes have been used for Araceae, Arecaceae, Commelinaceae, Poaceae, and Smilacaceae (Fig. 5). The monocotyledon node has been derived and is included in Figure 6. The result is different from that of Martin and Dowd (1989), and this is presumably because of the new computing program.

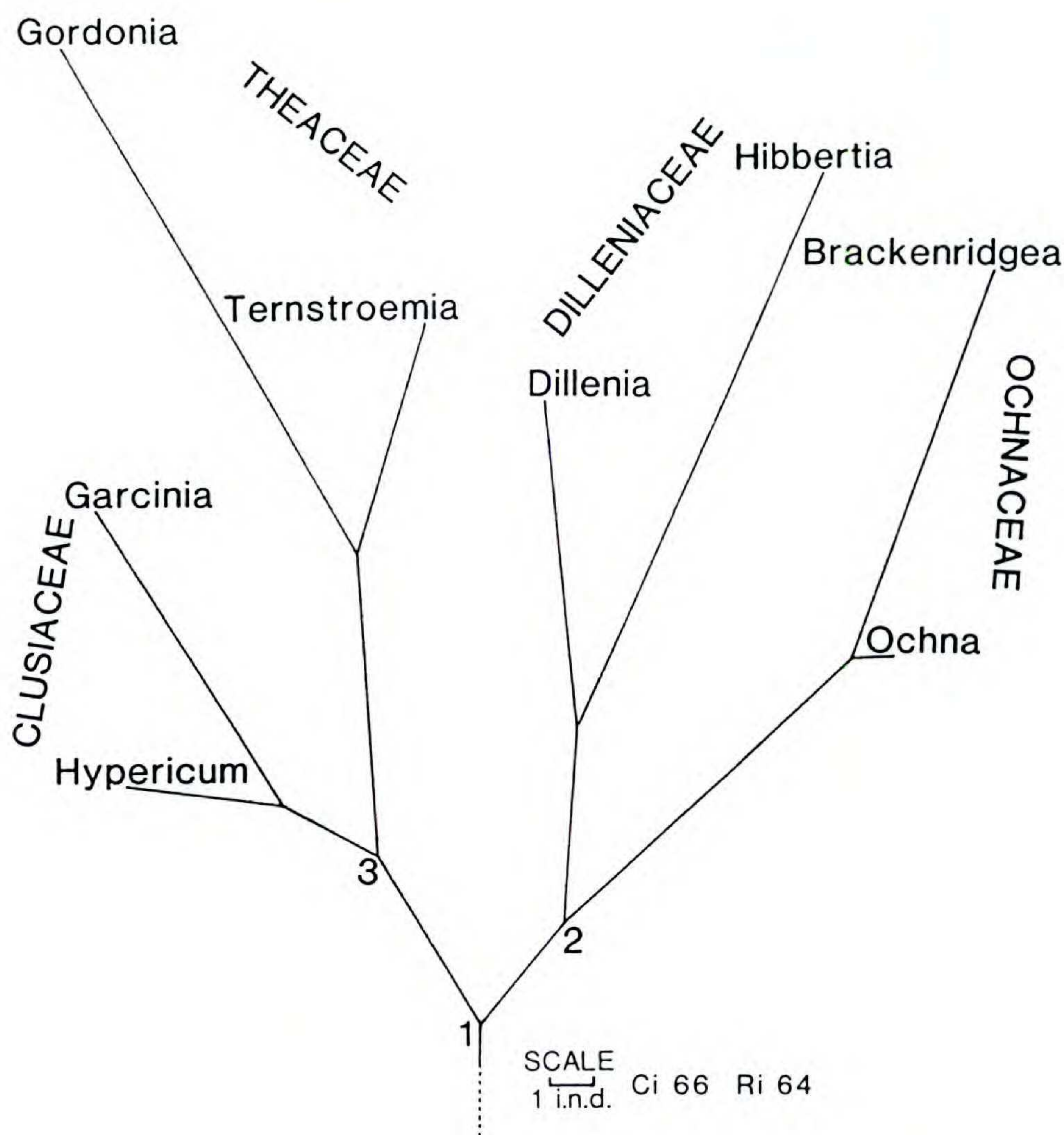


FIGURE 9. Group 6.

Group 4. There is good pairing between members of Ulmaceae and Urticaceae but not Moraceae (Fig. 7). It was only after failures with two species of *Ficus* and one of *Maclura*, almost certainly due to protease activity, that we supplemented with *Humulus*, knowing that its taxonomic position was not entirely clear. The failure of *Humulus* and *Morus* to pair was therefore not surprising. Subsequently, *Humulus* was removed from Group 4 and added to the list of uncertain taxa (see below).

Group 5. The tree of *Nothofagus* (Fig. 8b) is slightly different from that of Martin and Dowd (1988) because it is influenced by the weighting procedure of HENNIG86 and because it includes *Fagus*, which does not separate. From this tree a node has been derived and used in Figure 8a. While Betulaceae, Casuarinaceae, and Hamamelidaceae have correct grouping, the junction with Base divides *Nothofagus* from *Quercus*.

Group 6. Figure 9 differs from one already published by Martin and Dowd (1989); the two trees are of the same length but this one is preferred because it shows perfect pairing and reflects the

order when only familial nodes are used. However, the other probably conforms better with taxonomy in that Dilleniaceae is separate from the other three families.

Group 7. Figure 10 shows that the two representatives of Juglandaceae pair leaving *Myrica*, the sole representative of Myricaceae, separate.

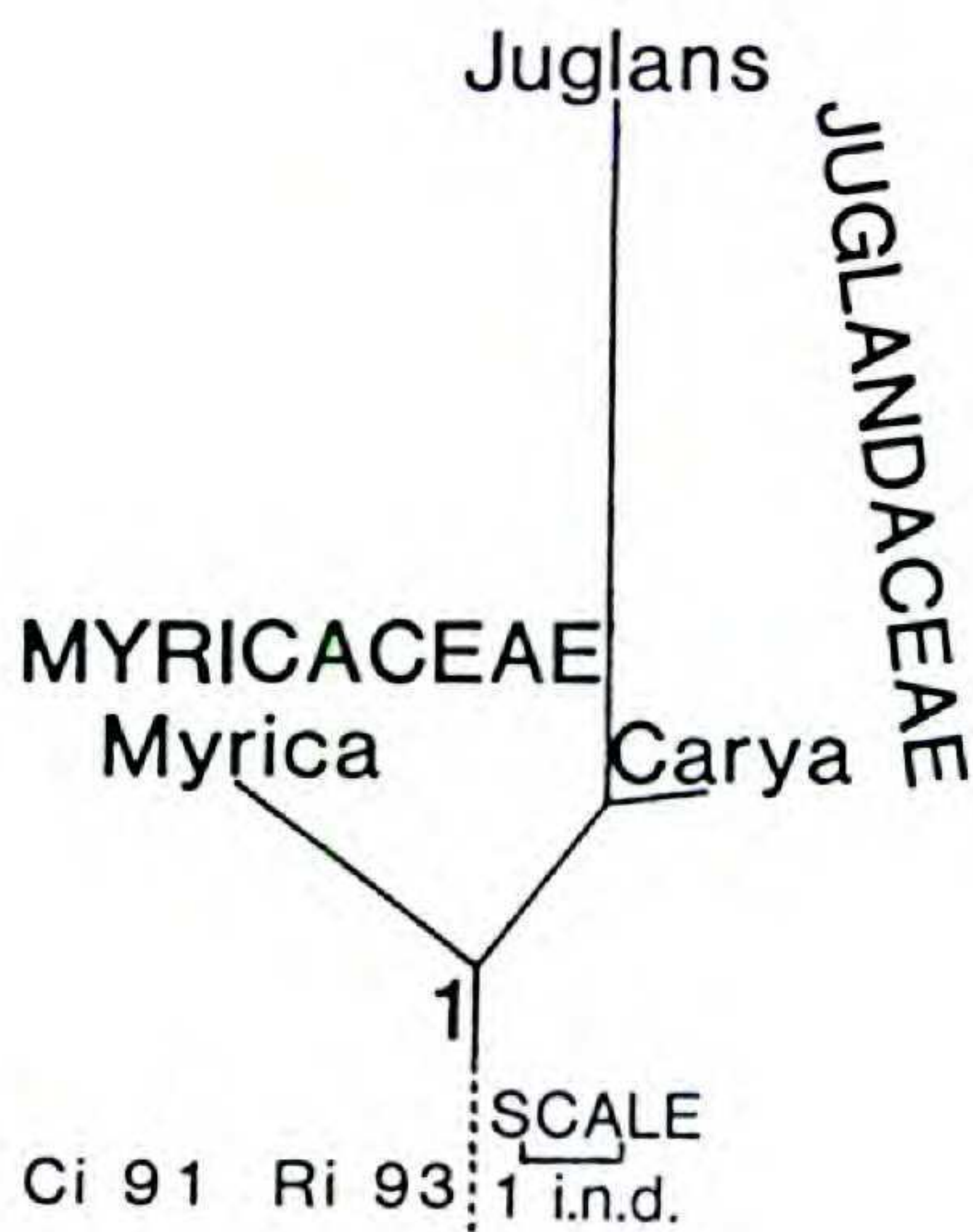


FIGURE 10. Group 7.

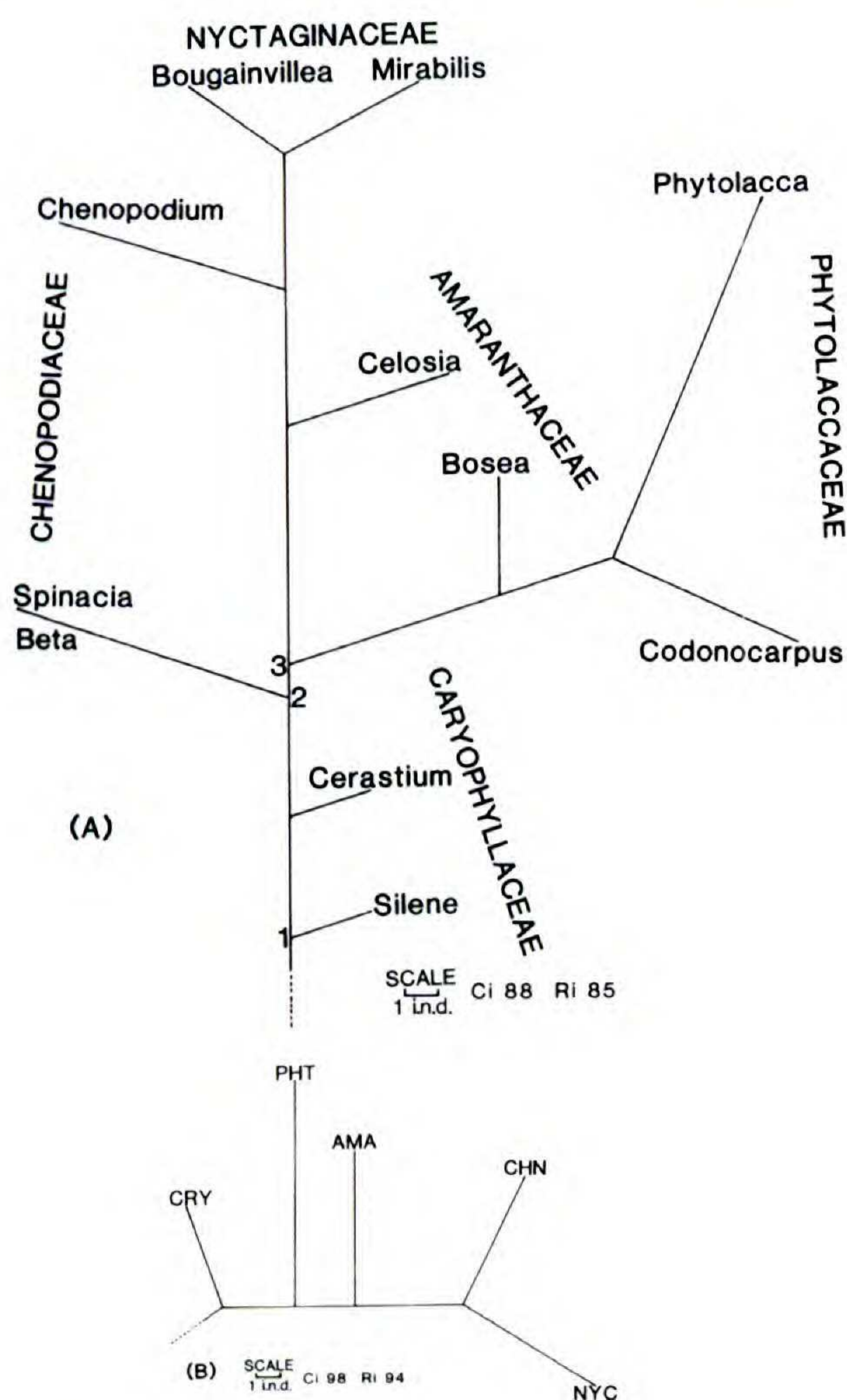


FIGURE 11.—(A). Group 8.—(B). Group 8 family nodes.

Group 8. The “Centrospermae” is one of the most unsatisfactory groups with representatives of two families, Chenopodiaceae and Amaranthaceae, failing to form pairs (Figure 11a). *Spinacia* and *Beta* have identical sequences, but these are quite different from *Chenopodium*. The tree for family nodes is in Figure 11b.

Group 9. There is a marked difference between the two representatives of Tiliaceae; *Grewia* is at the bottom of the tree (Fig. 12), while *Sparmannia* disrupts the clustering in Malvaceae. However, the remaining four families are satisfactory. *Grewia* was removed from Group 9 and added to the list of uncertain taxa. (As will be mentioned later, it subsequently rejoined.)

Group 10. While Violaceae, Cucurbitaceae, Salicaceae, Brassicaceae, and Flacourtiaceae formed good clusters (Fig. 13a), the two representatives of Datisceae (*Datisca* and *Tetrameles*) were very different. Attempts to study *Capparis* having failed, *Cleome* was left unpaired so we chose *Reseda* from the putatively related family Resedaceae. Since these two did group, we did not seek correct partners for them. In addition to these two

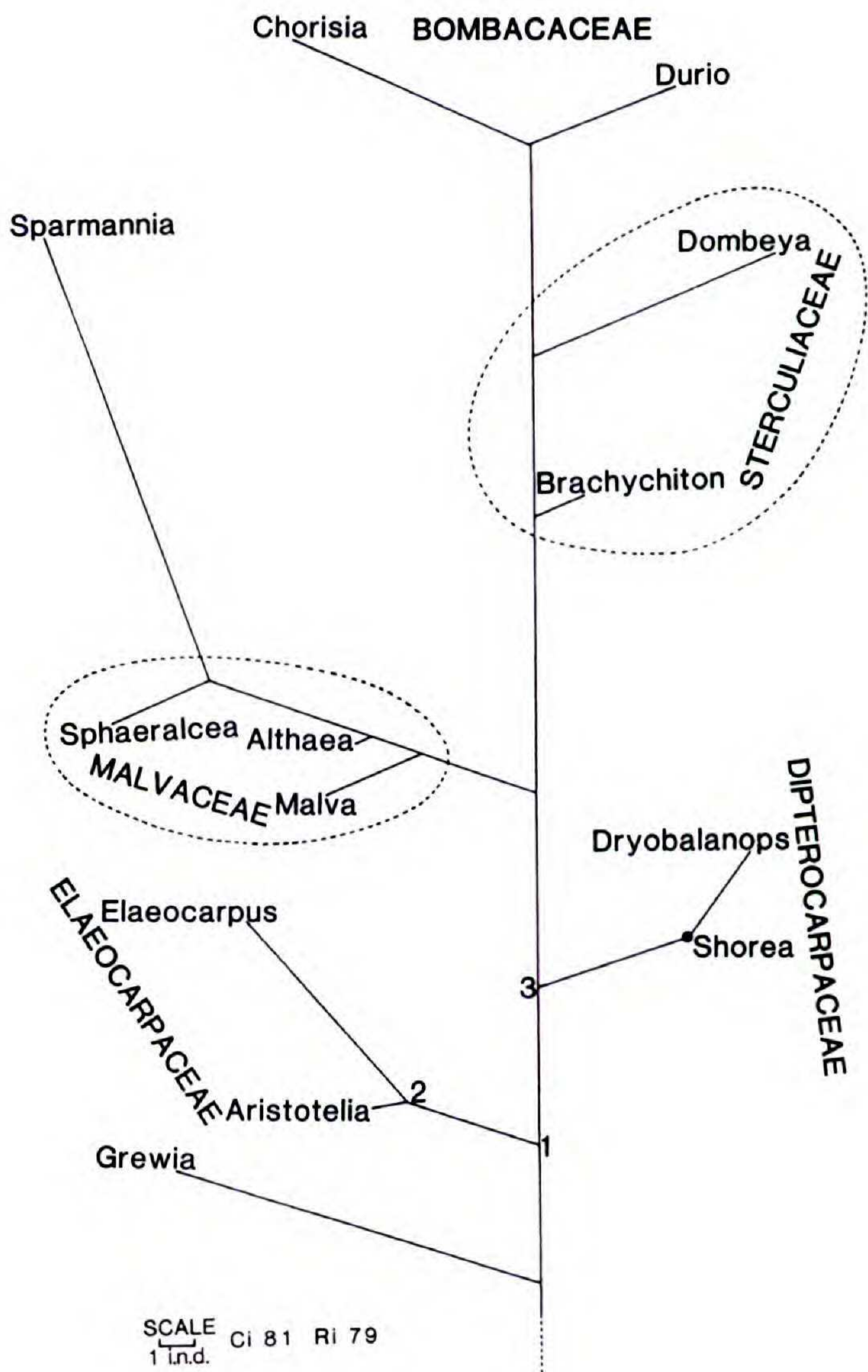


FIGURE 12. Group 9.

singletons, *Moringa* represents a monogeneric family. A tree from family nodes is shown in Figure 13b.

Group 11. The representatives of all four families form pairs (Fig. 14), Myrsinaceae adjacently and the other three families dichotomously.

Group 12. As mentioned earlier, the two species of Epacridaceae are distinguished by having an additional amino acid inserted in their sequences. Although this could have been used as a character, it was unnecessary because the two species paired separately from the two Ericaceae species (Fig. 15).

Group 13. When family nodes are derived for Rosaceae, Cunoniaceae, and Saxifragaceae, they are very close (Fig. 16b), so it is not surprising that there is confusion when individual species are analyzed (Fig. 16a). The representatives of Rosaceae pair correctly, however.

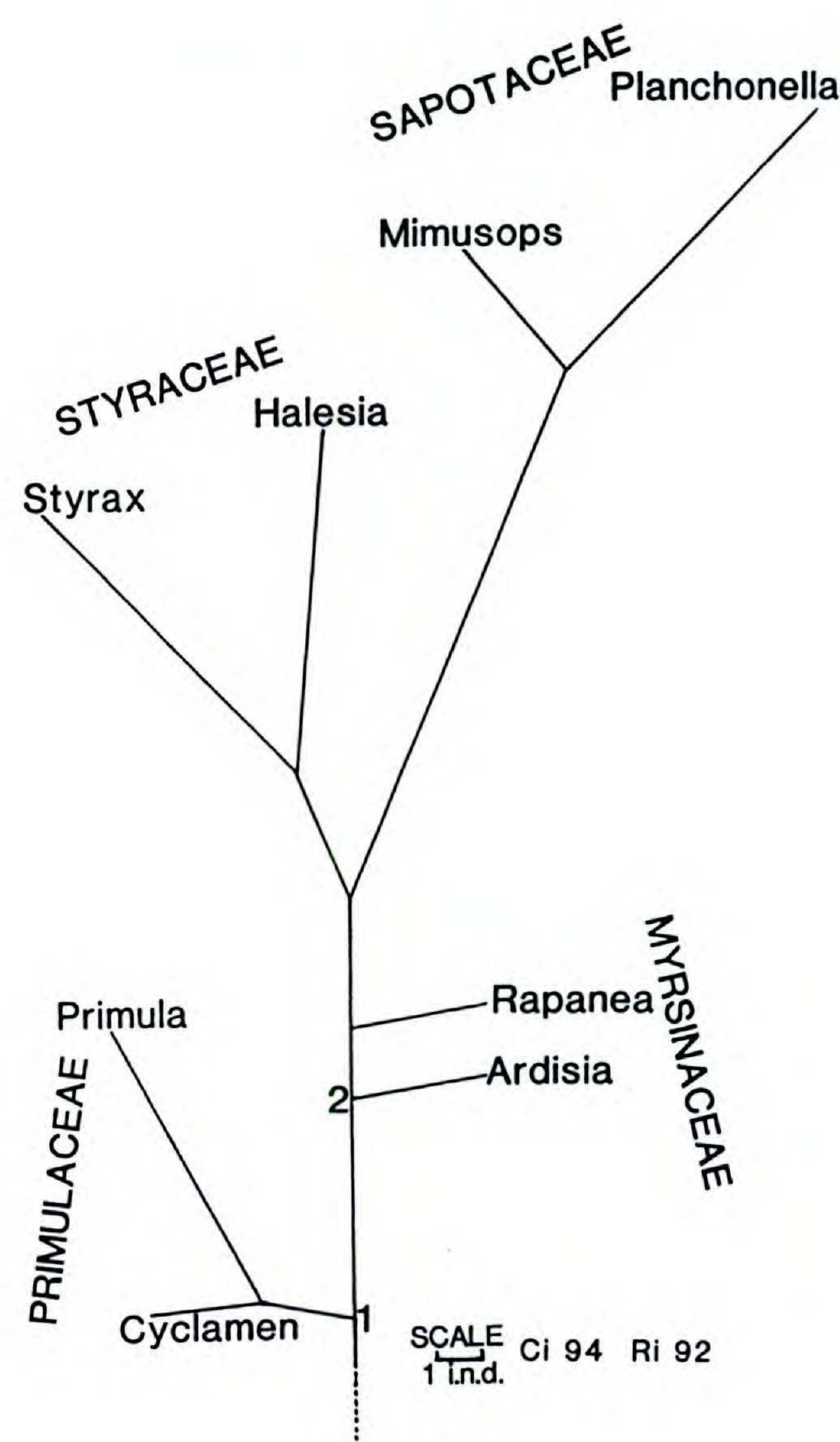
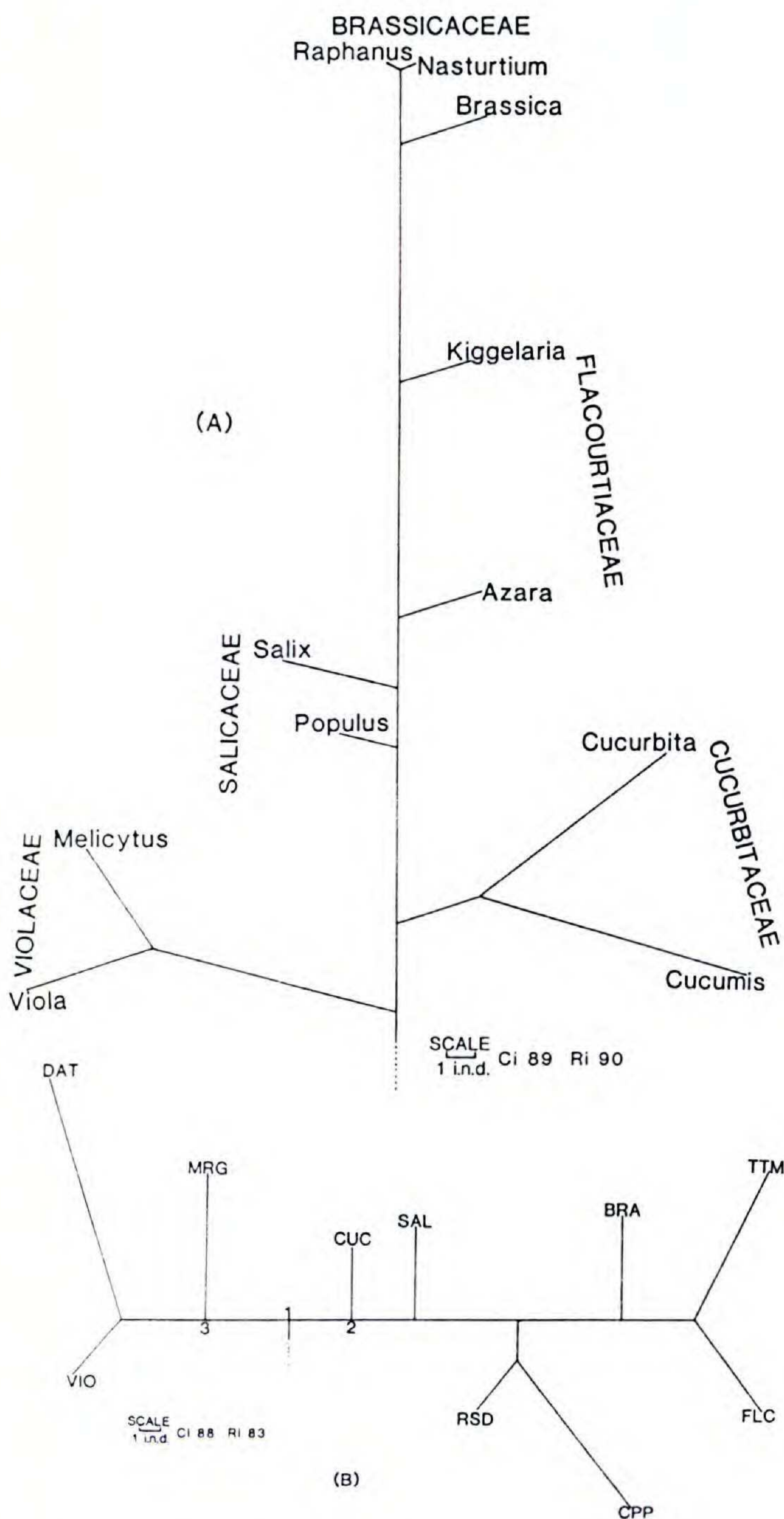


FIGURE 14. Group 11.

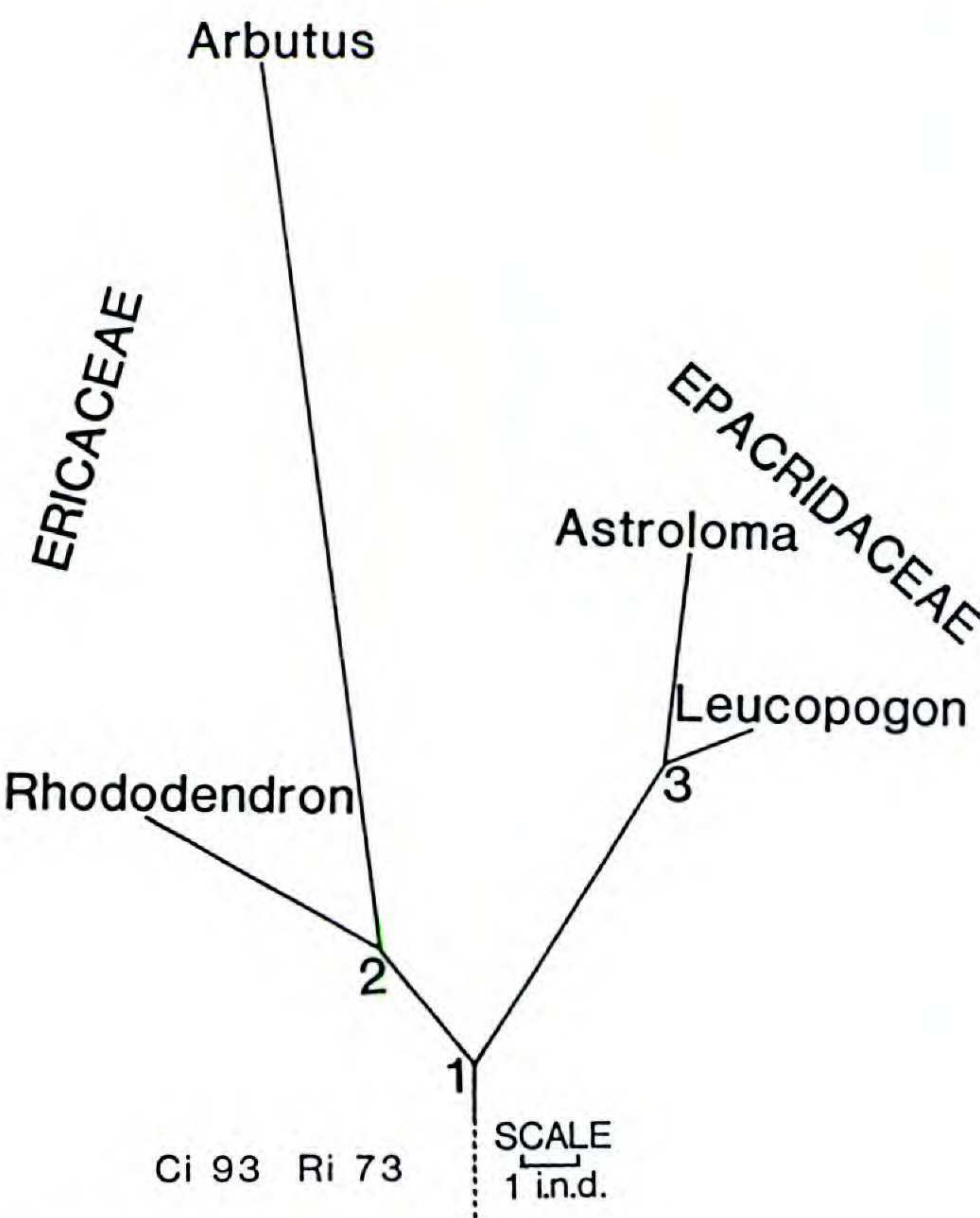


FIGURE 15. Group 12.

FIGURE 13. —(A). Group 10, omitting single representatives of families. —(B). Family nodes of Group 10. Because they did not pair and are sometimes placed in different families, *Datisca* and *Tetrameles* are included here.

Group 14. Among minimal trees derived when all legume species are analyzed simultaneously, there are some in which the two Mimosaceae species pair and so do the two Caesalpiniaceae; however, the eight Papilionaceae species are confused. We have therefore derived a Papilionaceae node separately (Fig. 17b) and show this with the other two families (Fig. 17a).

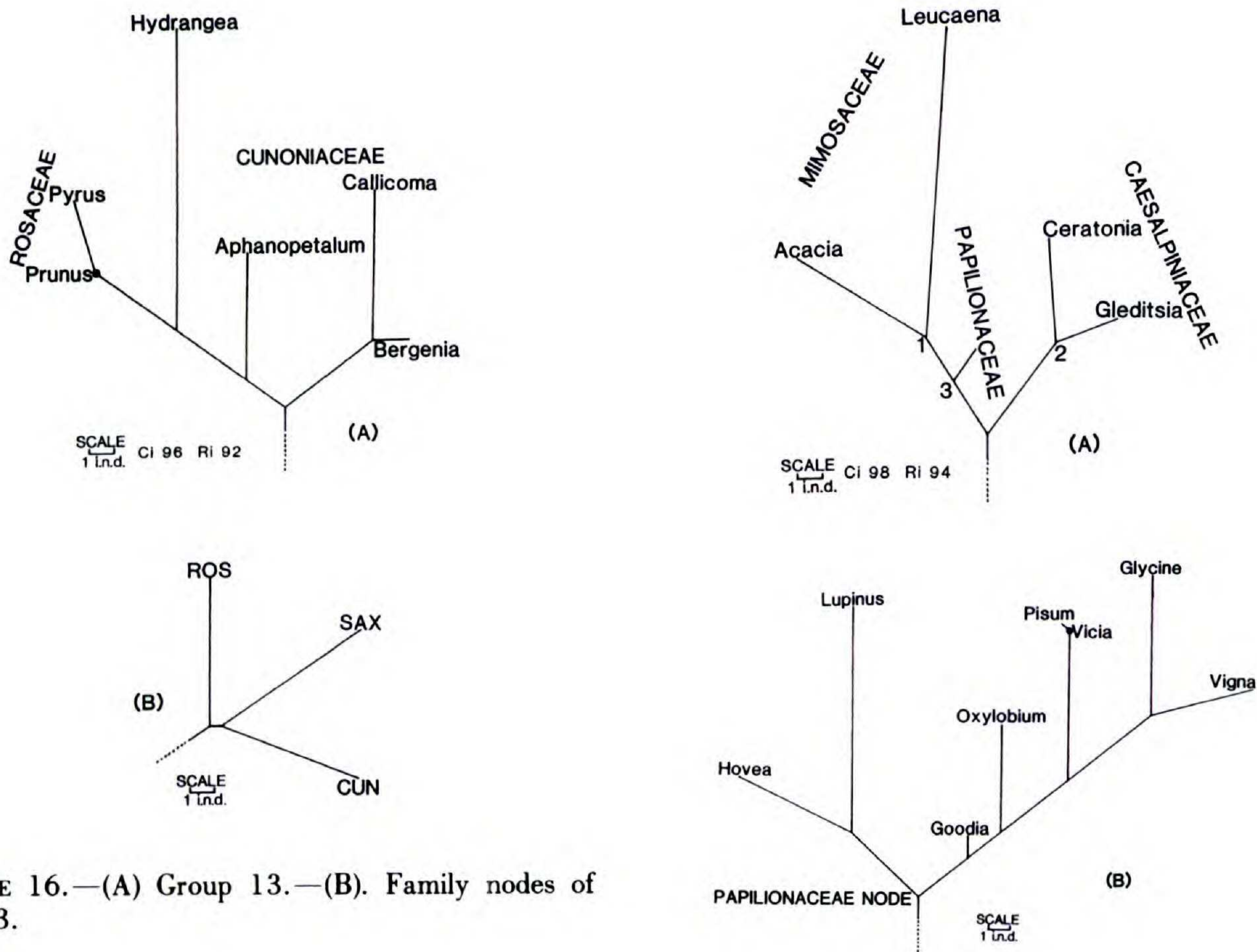


FIGURE 16.—(A) Group 13.—(B). Family nodes of Group 13.

FIGURE 17. (A) Group 14 with Papilionaceae represented by a node derived from (B).

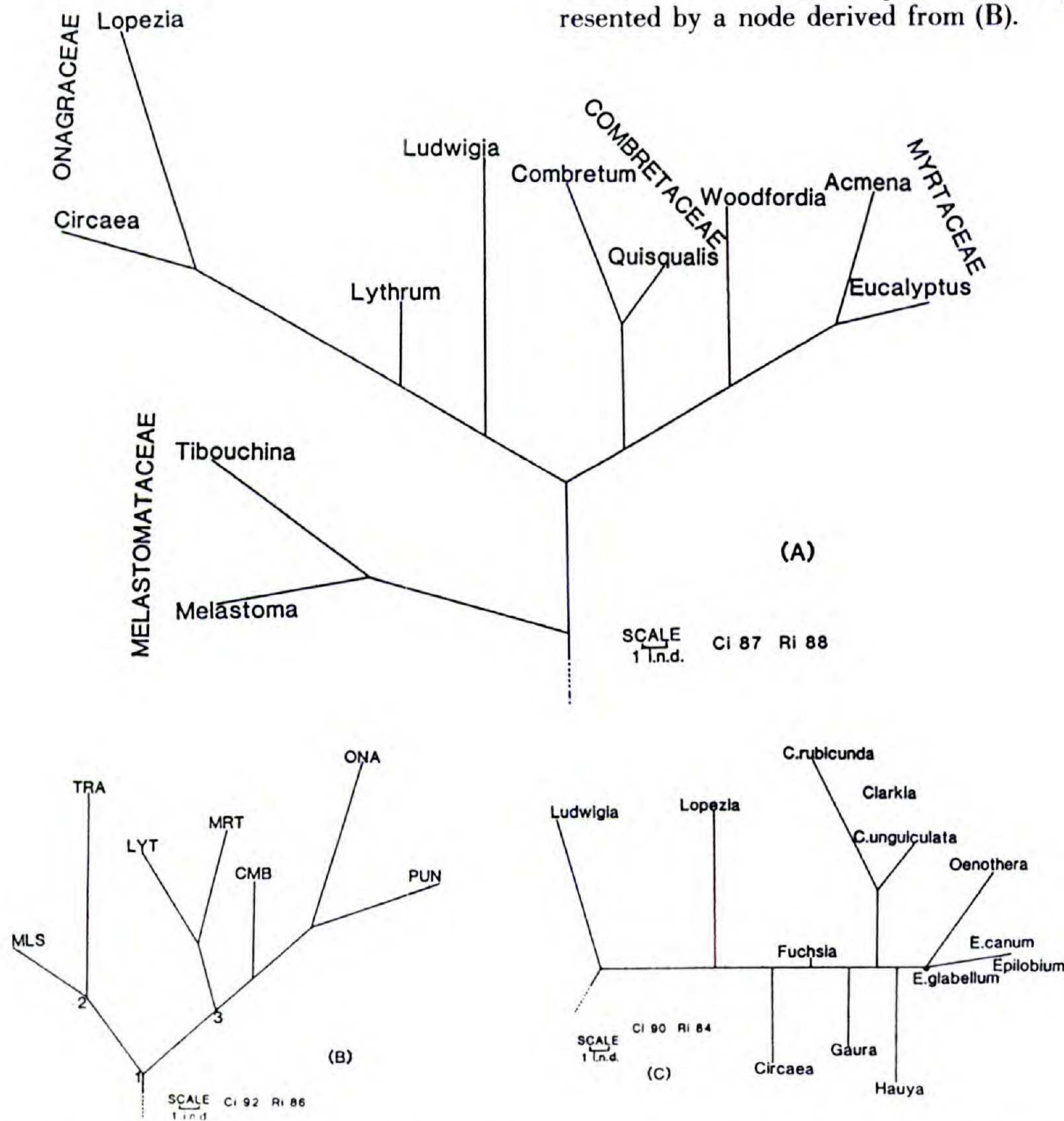


FIGURE 18. (A) Group 15 omitting *Trapa* and *Punica*, which are included with family nodes in (B). (C) Onagraceae.

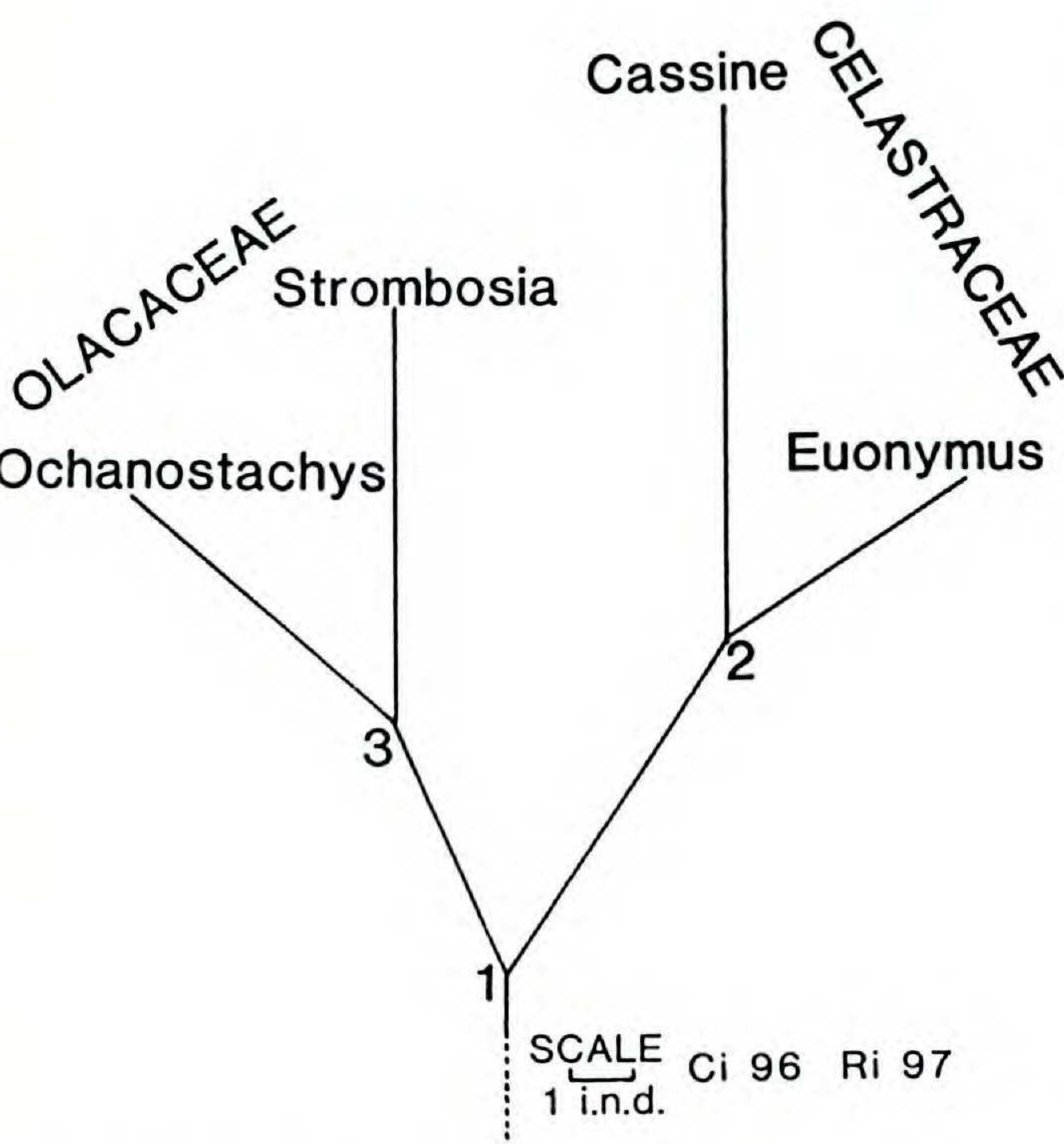


FIGURE 19. Group 16.

Group 15. This Group, which corresponds to the order Myrtales, was discussed by Martin and Dowd (1986a). Since then only *Trapa* and *Punica*, both singletons, have been added. When the representatives of the other five families are analyzed, pairing is good except in Lythraceae (Fig. 18a). The three members of Onagraceae in this tree are from the bottom of the family tree (Fig. 18c). When family nodes are analyzed (Fig. 18b), the root of the tree is in a different place from the one previously published; it is uncertain why this is so, but

it could be due to the inclusion of new families, the earlier choice of inappropriate outgroups, or the new analytical methods.

Group 16. As discussed earlier, all representatives of the hemiparasites of the Santalales and Loranthaceae failed to yield protein samples, so this Group is reduced to Olacaceae and Celastraceae in which pairings are straightforward (Fig. 19).

Group 17. This Group is not very satisfactory possibly because, as indicated in Figure 20b, there has been a rapid radiation. The consequences are that the members of Simaroubaceae and Sapindaceae do not pair, while *Flindersia*, sometimes excluded from Rutaceae, does not group with the other two representatives of that family. However, there is good pairing for Connaraceae and Anacardiaceae (Fig. 20a). *Melia* having failed, *Cedrela* is left as the sole representative of Meliaceae.

Group 18. The two members of Haloragaceae, *Gonocarpus* and *Haloragodendron*, are so confounded with the three members of Rhizophoraceae (Fig. 21) that there was no point in deriving family nodes to derive a Group node.

Group 19. *Nitraria* having failed, *Zygothymum* is a singleton as is *Tropaeolum*, for which no partner was available. As shown in Figure 22a, the members of Geraniaceae and Malpighiaceae pair. The family node tree is shown in Figure 22b.

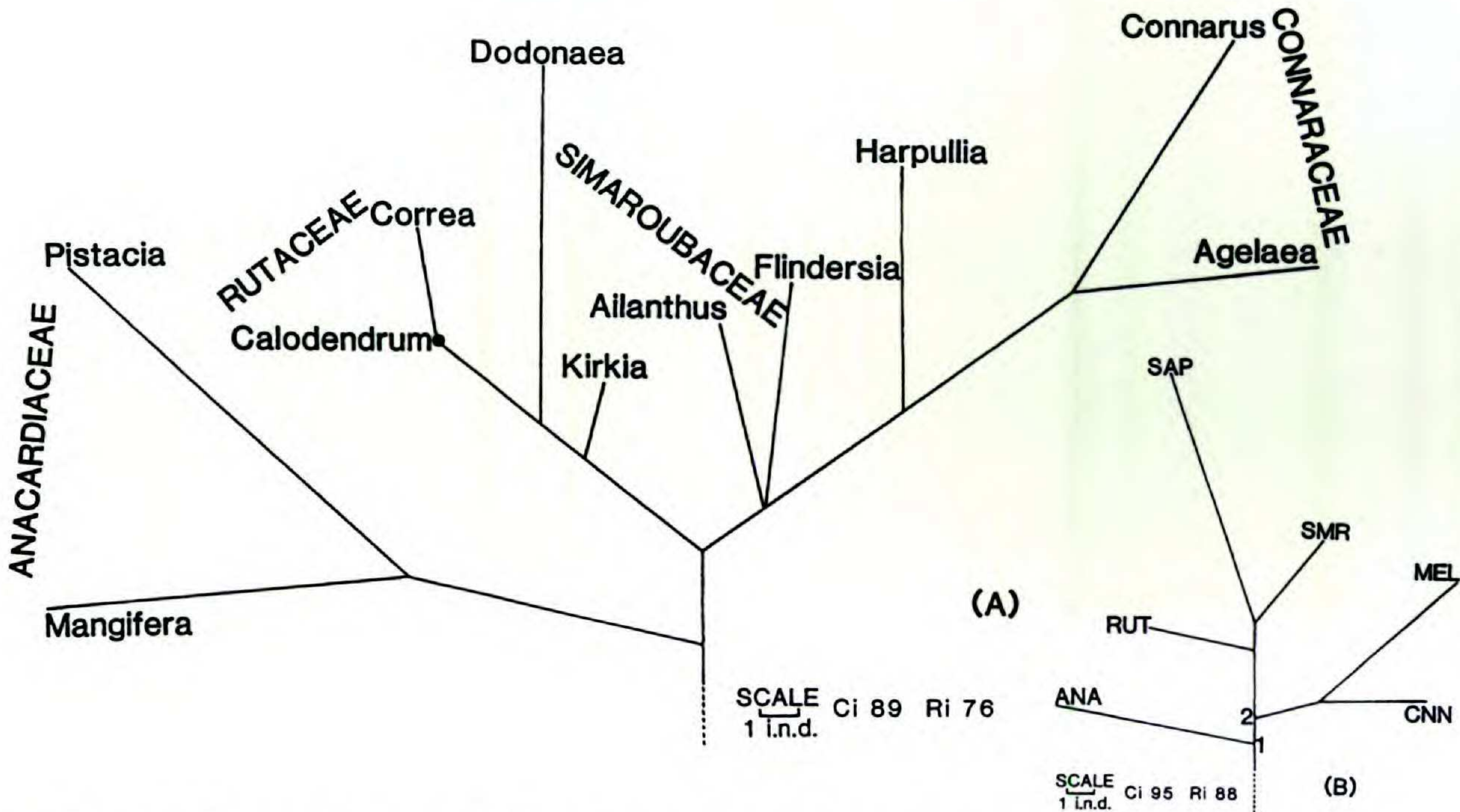


FIGURE 20. (A) Group 17 omitting *Melia*, which is included with family nodes in (B).

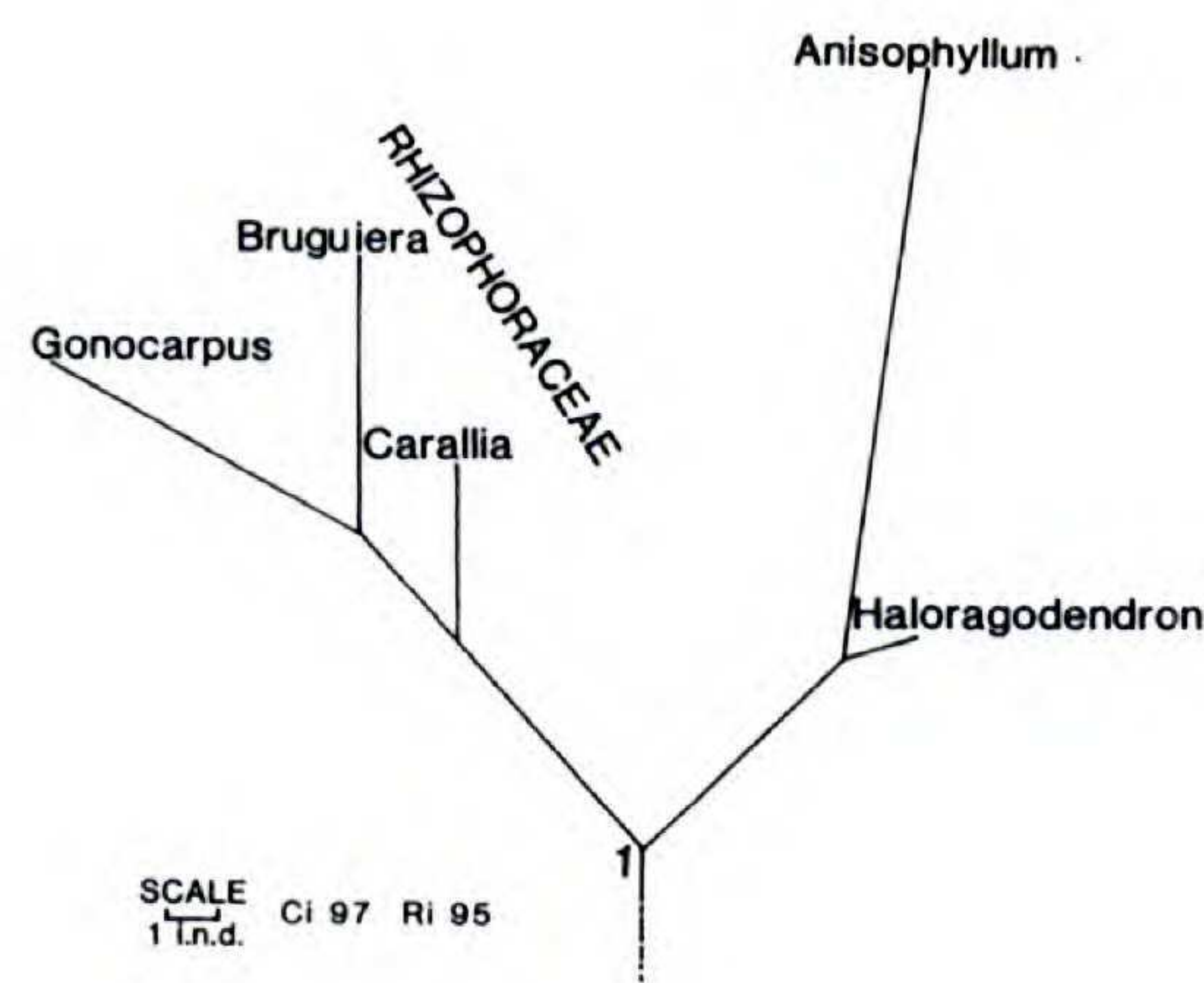


FIGURE 21. Group 18.

Group 20. *Hoya* was left a singleton by failure to extract protein from two other members of Asclepiadaceae, *Asclepias* and *Cryptostegia*. The members of four families showed dichotomous pairing while *Logania* paired alongside *Buddleia*, which is only possibly a member of Loganiaceae (Fig. 23).

Group 21. The representatives of Lamiaceae and Verbenaceae were very similar but there was nevertheless a minimal tree in which congruent pairing occurred (Fig. 24).

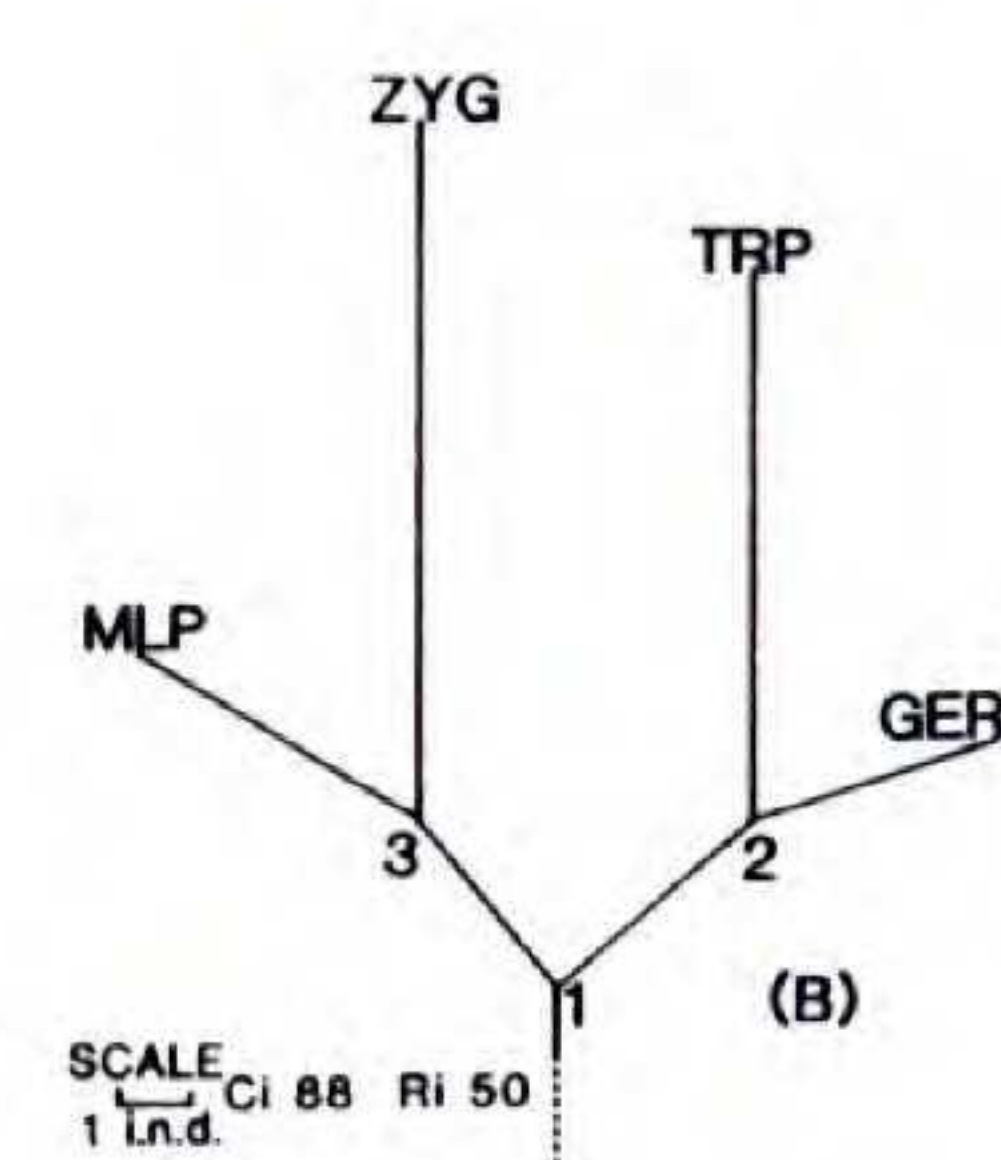
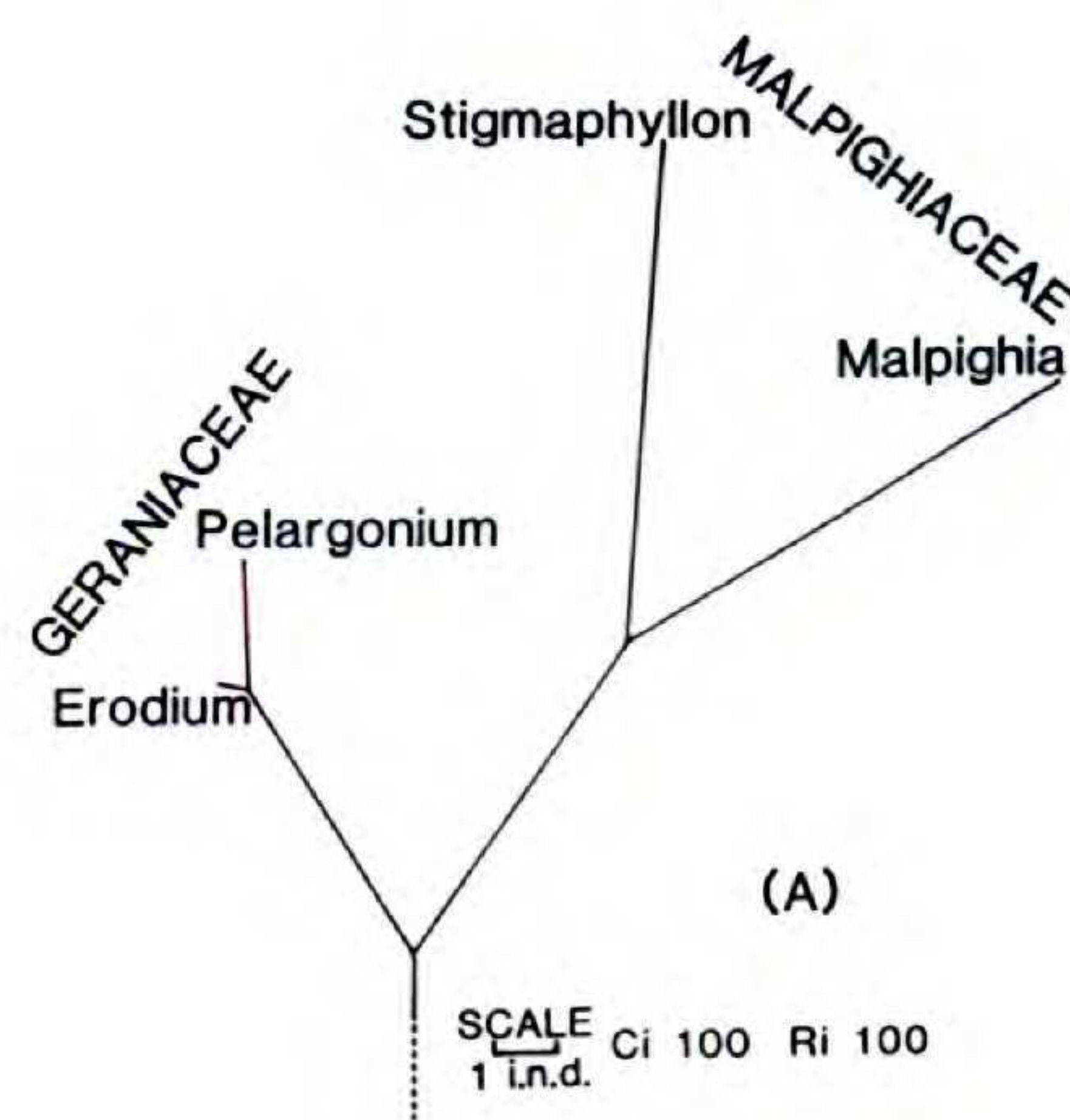


FIGURE 22. (A) Group 19 omitting *Zygophyllum* and *Tropaeolum*, which are included in (B).

Group 22. There have been previous reports of *Solanum* (Martin et al., 1986) and *Nicotiana* (Martin & Dowd, 1984b). Using HENNIG86, new

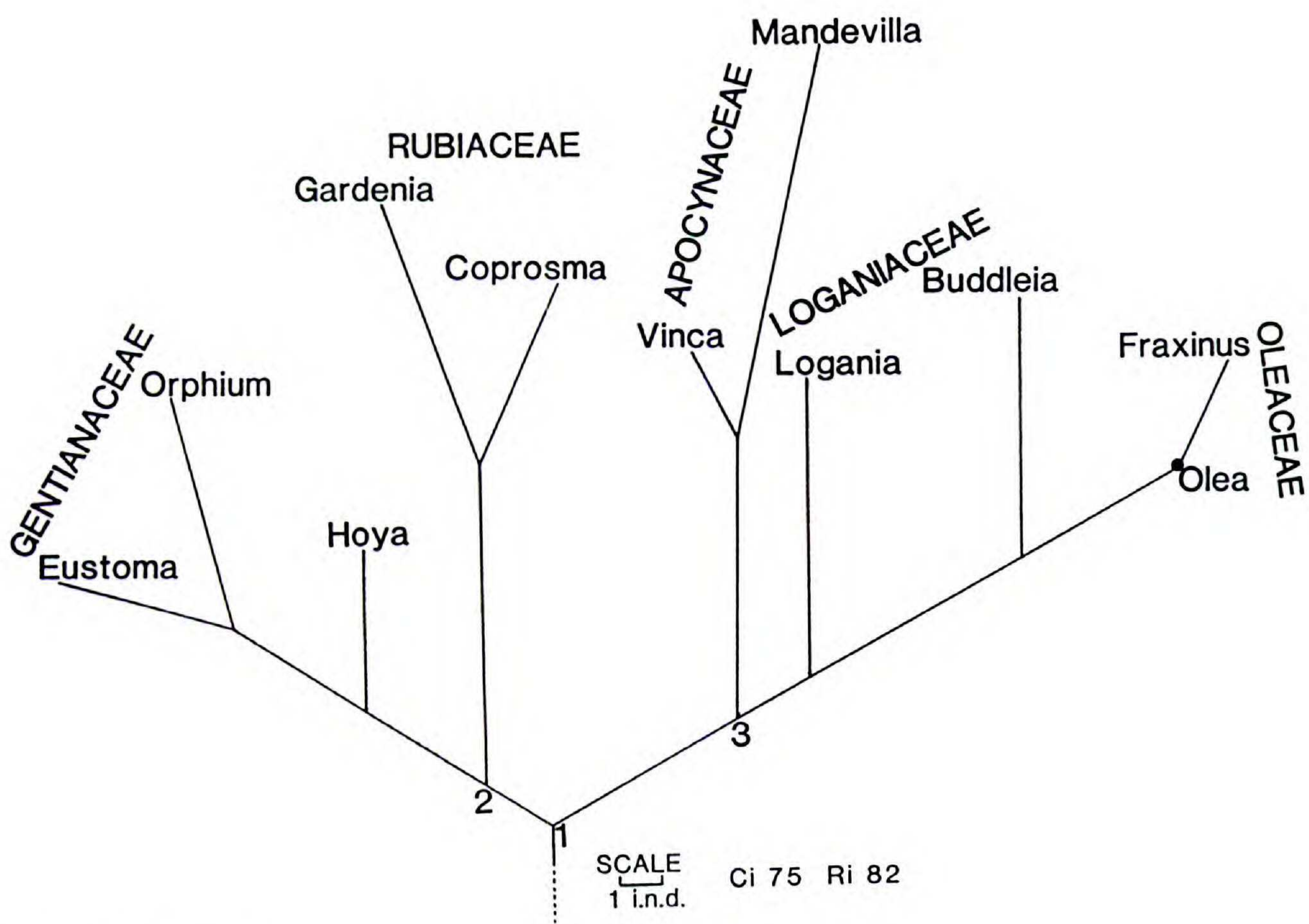


FIGURE 23. Group 20.

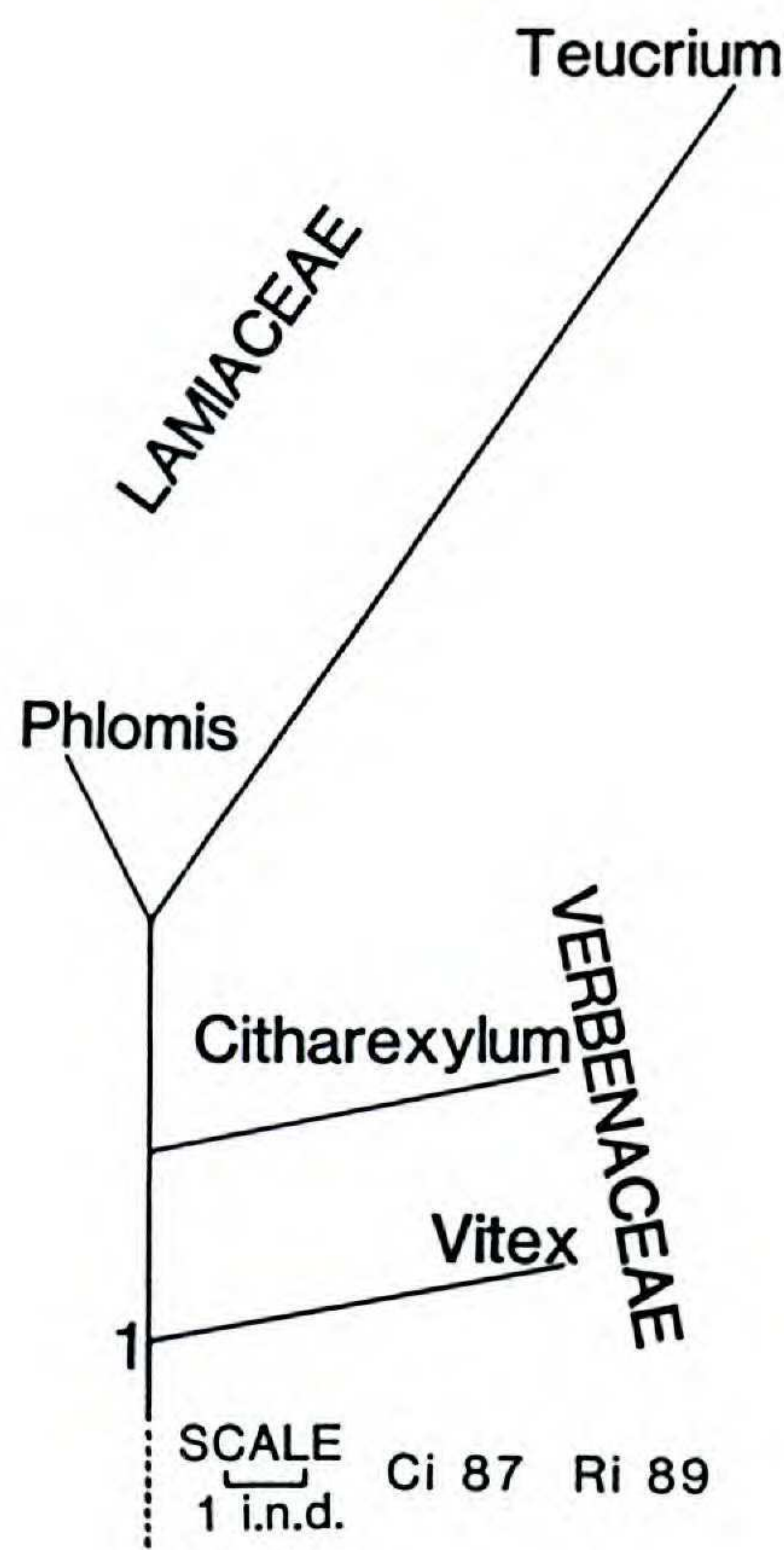


FIGURE 24. Group 21.

nodes have been derived for both genera (Fig. 25b, c) and were used, with *Anthocercis*, to represent Solanaceae in Figure 25a. These group well but there is confusion between the representatives of Convolvulaceae and Polemoniaceae.

Group 23. As mentioned earlier, all attempts to extract rubisco from Acanthaceae species failed. The representatives of the other four families of this Group pair well, Scrophulariaceae, Gesneriaceae, and Bignoniaceae dichotomously and Pedaliaceae adjacently (Fig. 26).

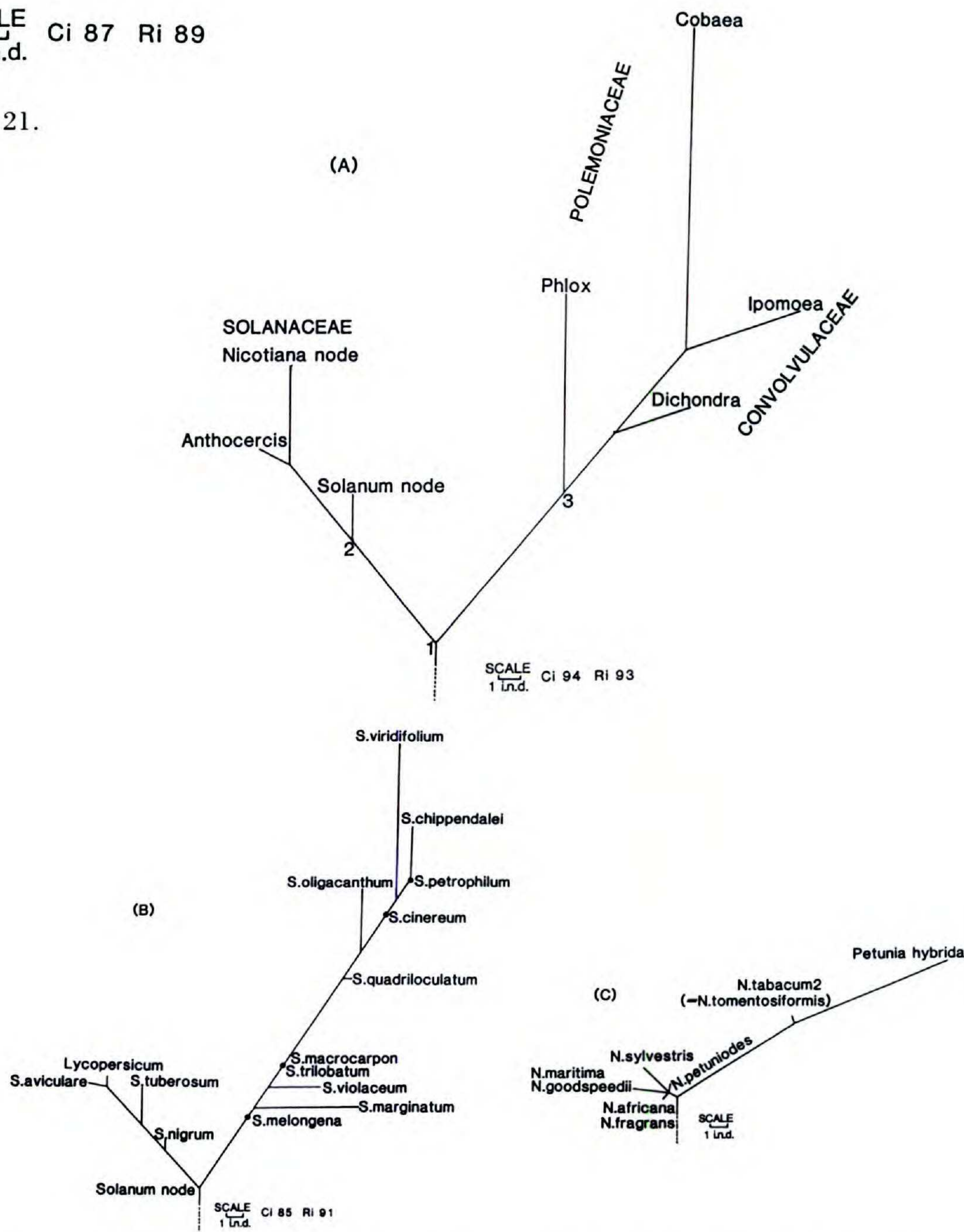


FIGURE 25. (A) Group 22 with the *Solanum* node derived from (B) and the *Nicotiana* node from (C).

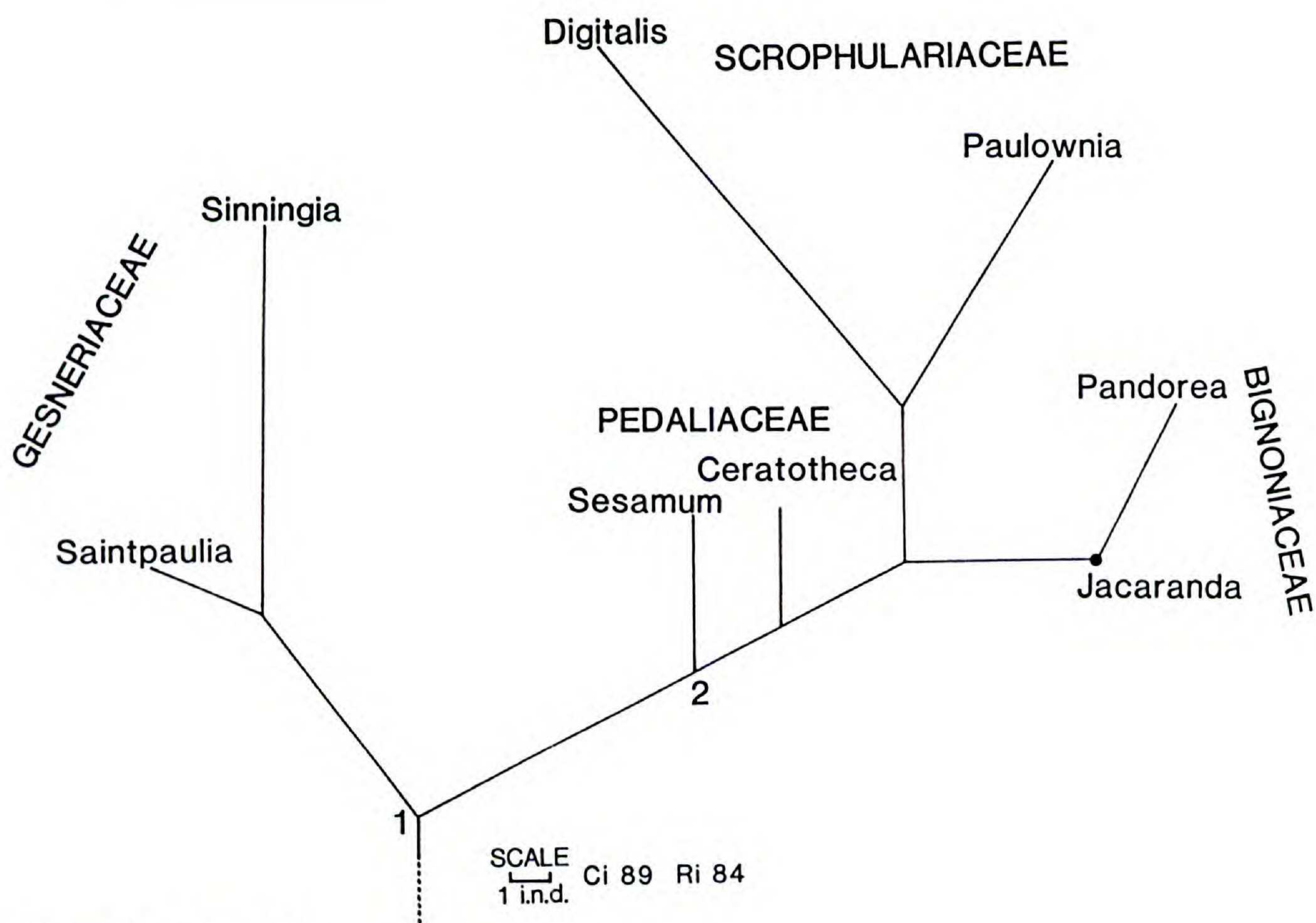


FIGURE 26. Group 23.

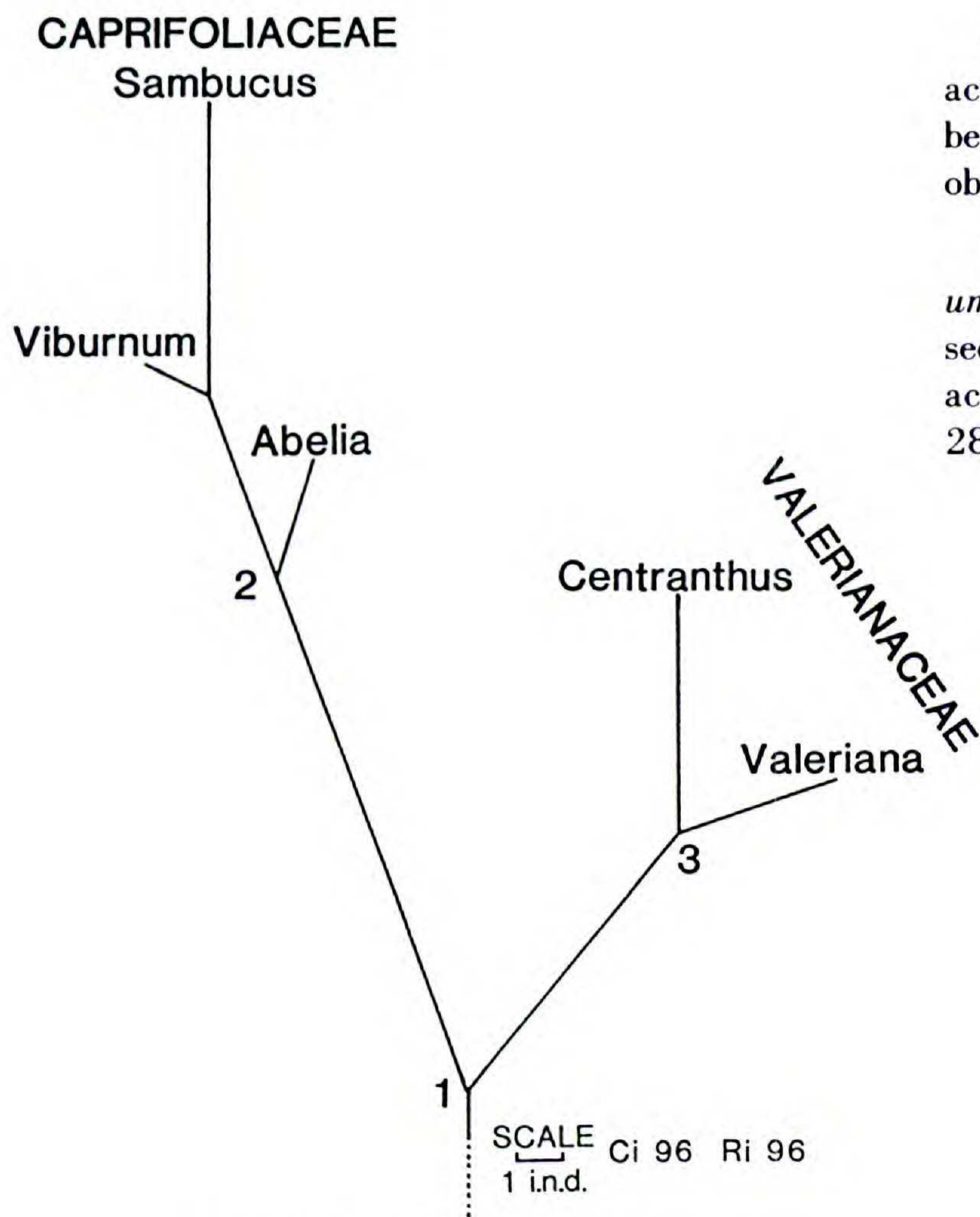


FIGURE 27. Group 24.

Group 24. The three members of Caprifoliaceae are substantially different from the two members of Valerianaceae so that correct grouping is observed (Fig. 27).

Group 25. This Group is unusual in that *Apium* and *Foeniculum* of Apiaceae have identical sequences as do *Schefflera* and *Fatsia* of Araliaceae. Consequently, the tree of this Group (Fig. 28) is very simple.

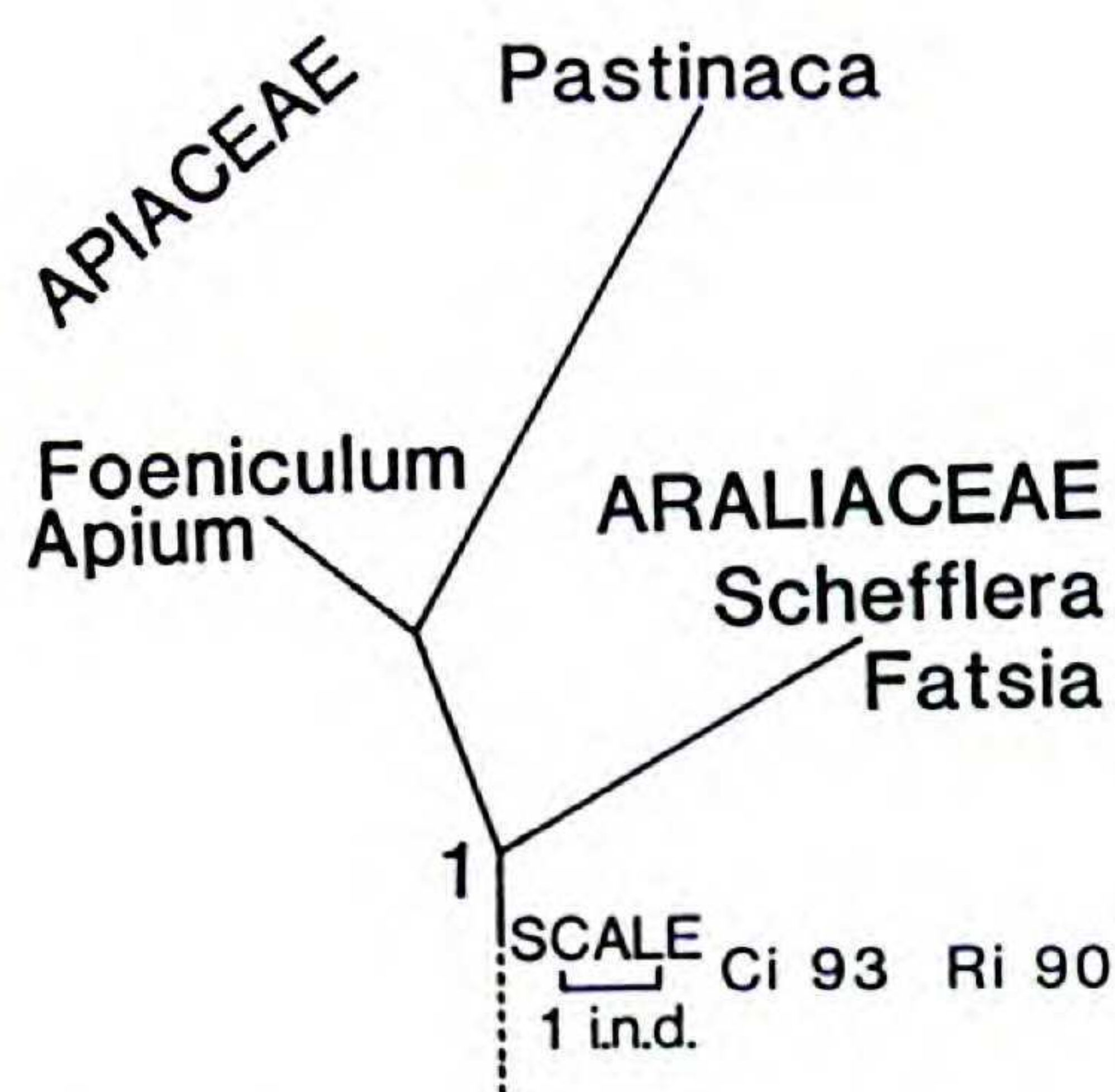


FIGURE 28. Group 25.

THE DERIVATION OF A TREE FOR THE GROUPS OF DICOTYLEDONS

FIRST STAGE; A TEST OF THE REALITY OF GROUPS

Depending on the size and complexity of the Group, one, two, or three nodes have been marked near the bases of each Group tree; altogether there are 58 basal nodes and the ancestral sequence of each has been derived using ANALYZE. These have been used for a test of the reality or integrity of the Groups. If a family does not really belong to a Group, it should usually behave like an out-group and assume the position closest to the base of the tree. Thus, in a simultaneous analysis of all 58 basal nodes, it would be expected that nodes truly belonging to the same Group should cluster together. If a family is misplaced in a Group, the nodes should separate.

The only program that can be used with 58 taxa simultaneously is HENNIG86 with the option *mhennig* followed by *bb*. This was done three times, each yielding large numbers of trees for which strict consensus trees were derived. Inspection indicated that most Groups behaved as if they were real, but some separation of nodes occurred in Groups 5, 8, 14, 15, 22, and 24 (see below). It is unlikely that this sort of analysis would give a completely reliable result, but our interpretation is that where there is no separation of within-Group nodes, that Group should be accepted as valid. We understand that our test is not infallible, but we are reluctant, at this stage, to attempt another obvious test, viz. the simultaneous analysis of adjoining Groups. This test was used earlier with Groups 1, 2, and 3 and led to considerable mixing of the first two. The amount of convergent evolution between Groups is probably such that, if this test were applied widely, confusion would result. Therefore, even though we understand the limitations, we confine our testing of the integrity of Groups to one sort of analysis.

For the six Groups where there was doubt, we applied the test devised by Lake (1987). This is confined to four species, A, B, C, D and uses a chi-square test to decide which is the most probable of the three possible relationships, viz. A + B & C + D or A + C & B + D or A + D & B + C. Thus representatives of each part of a divided Group were tested with representatives of the Groups with which they most closely clustered. These tests gave no further grounds for doubting the integrity of Groups 8 and 15 and consequently, in the next stage of the analysis, they were included un-

changed. The tests reinforced the doubts about Groups 14, 22, and 24, so their separate parts were added to the list of uncertain families to be incorporated later. These were: from Group 14, Mimosaceae plus Papilionaceae on the one hand and Caesalpiniaceae on the other; from Group 22, Convolvulaceae plus Polemoniaceae on the one hand and Solanaceae on the other; from Group 24, Valerianaceae and Caprifoliaceae. Tests with Group 5 were equivocal, so Hamamelidaceae was removed and added to the list of uncertain families, but the node for the remaining three families was used at the next stage.

SECOND STAGE; DERIVING A PRELIMINARY, ABBREVIATED TREE

Following the first stage, the basal node was used to represent each of the 22 remaining Groups (though amended in Group 5 after removal of Hamamelidaceae). Several analyses, using *mhennig* and *bb*, were carried out on these nodes. The object was to identify apparently constant associations from which nodes might be derived in order to reduce the number of taxa to 16, a number compatible with analyses using the reliable *ie* program at the next stage. The following five pairings were chosen and their nodes derived: Groups 2 and 3; Groups 6 and 16; Groups 7 and 9; Groups 11 and 19; Groups 21 and 23. Group 25 was omitted at this stage because it was small, well-defined by morphology and our own work and could be incorporated later in the same way as the uncertain taxa. The resulting tree of 16 taxa is shown in Figure 29.

THIRD STAGE; INCORPORATING TAXA OF UNCERTAIN AFFINITIES

At this point there were 28 taxa on the list of those with uncertain affinities, comprising 18 families that were not placed in a Group (see Table 1 but note that Piperaceae and Nelumbonaceae were considered earlier), two genera (*Humulus* and *Grewia*) excluded from Groups during their analysis, five families and two pairs of families excluded during the first stage, and Group 25 omitted at the second stage. We wanted to add these into the second stage tree as accurately as possible using the *ie* program. These analyses with 17 taxa could each be performed in about a day. Although there was some variation, the second stage tree remained reasonably stable during these analyses, and we noted where each uncertain taxon fit. Six joined

in the basal third, 14 in the middle third, and 16 in the distal third. (The nonadditivity reflects that rigid demarcation was not exercised and borderline taxa were placed in two sets.) The members of each of the three sets were then analyzed with the corresponding members of the second stage tree and possible new or amended Groups were identified.

FOURTH STAGE; REDEFINITION OF GROUPS

Putative new or amended Groups were tested extensively to ensure that they were real. In this process an important factor in determining the coherence of Groups was the length of the internode joining a hitherto uncertain member to the Group. Penny et al. (1987) have emphasized that “long edges attract,” and we have long been aware that the junction of a distantly connected taxon is subject to so much variation that it is scarcely reliable. Thus, we have usually rejected a potential new member of a Group if it joins with a disproportionately long internode and have left it as uncertain.

In three cases, definite hypotheses arising from our work could be tested. In each case the question was whether a taxon belonged to the Group to which it was initially assigned (Table 1) or to the Group indicated by stage 3. This could be answered by considering the lengths of the alternative trees. In two cases, *Humulus* and Hamamelidaceae, the new grouping was shorter and therefore preferred. For *Grewia*, the trees were the same length so there was no good reason for preferring the new grouping (with Group 18).

As a consequence of these tests, only 15 of the original 25 Groups have the same composition as they had before the first stage of this section. The other ten Groups have been increased, decreased, or merged. Where a nucleus of an original Group remains, the number has been retained but A added. Original Groups 13, 24, and 25 have disappeared. New Groups 26, 27, 28, and 29 have been formed.

Group 4A. *Humulus* has been removed.

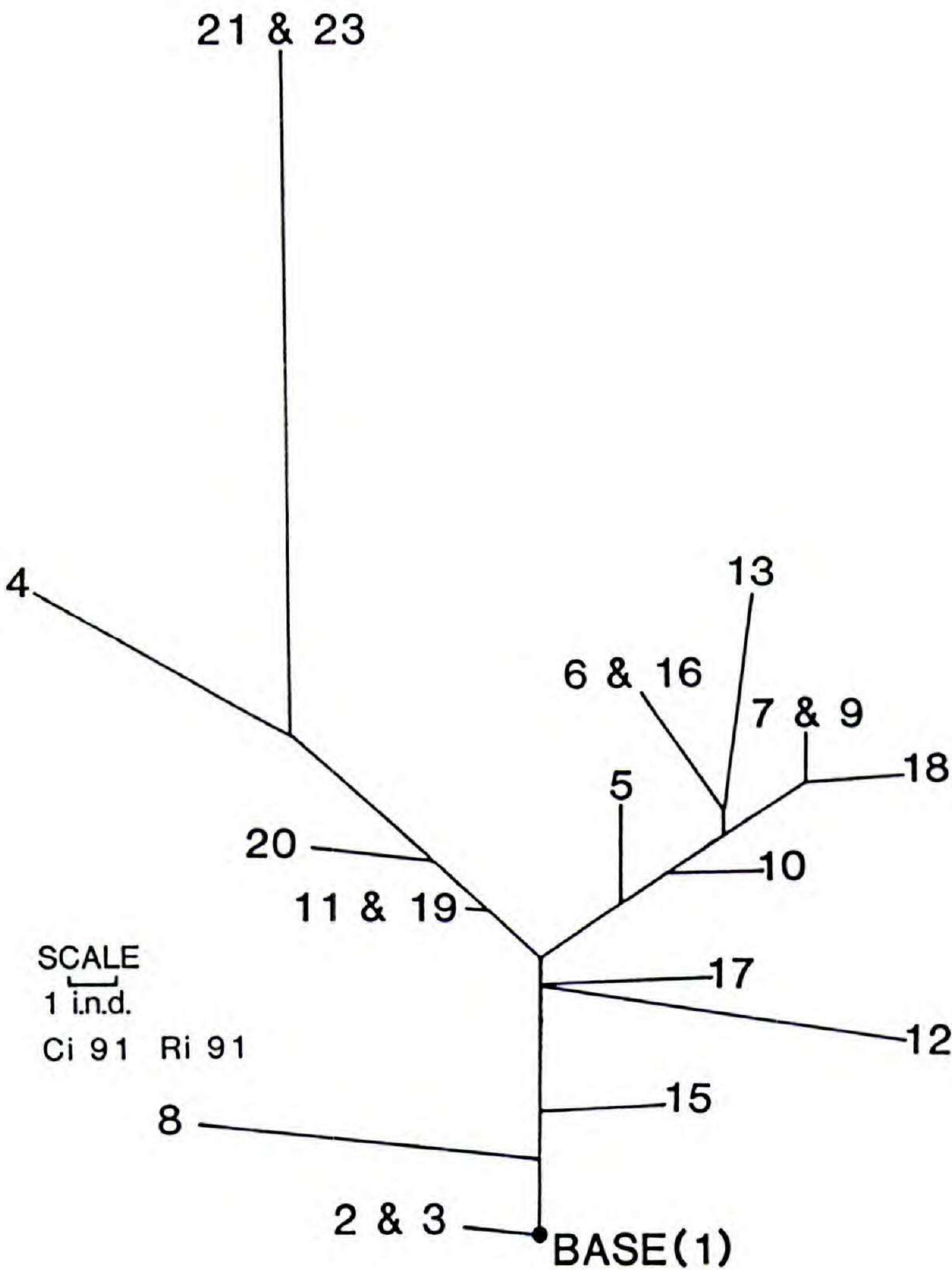


FIGURE 29. The provisional tree of Group nodes abbreviated by combining some Groups and omitting others that had dissolved.

Group 5A. Hamamelidaceae has been removed.

Group 8A. Although the original Group 8 (“Centrospermae”) remains intact, Lecithydaceae and *Humulus* join the same branch of the tree (Fig. 30).

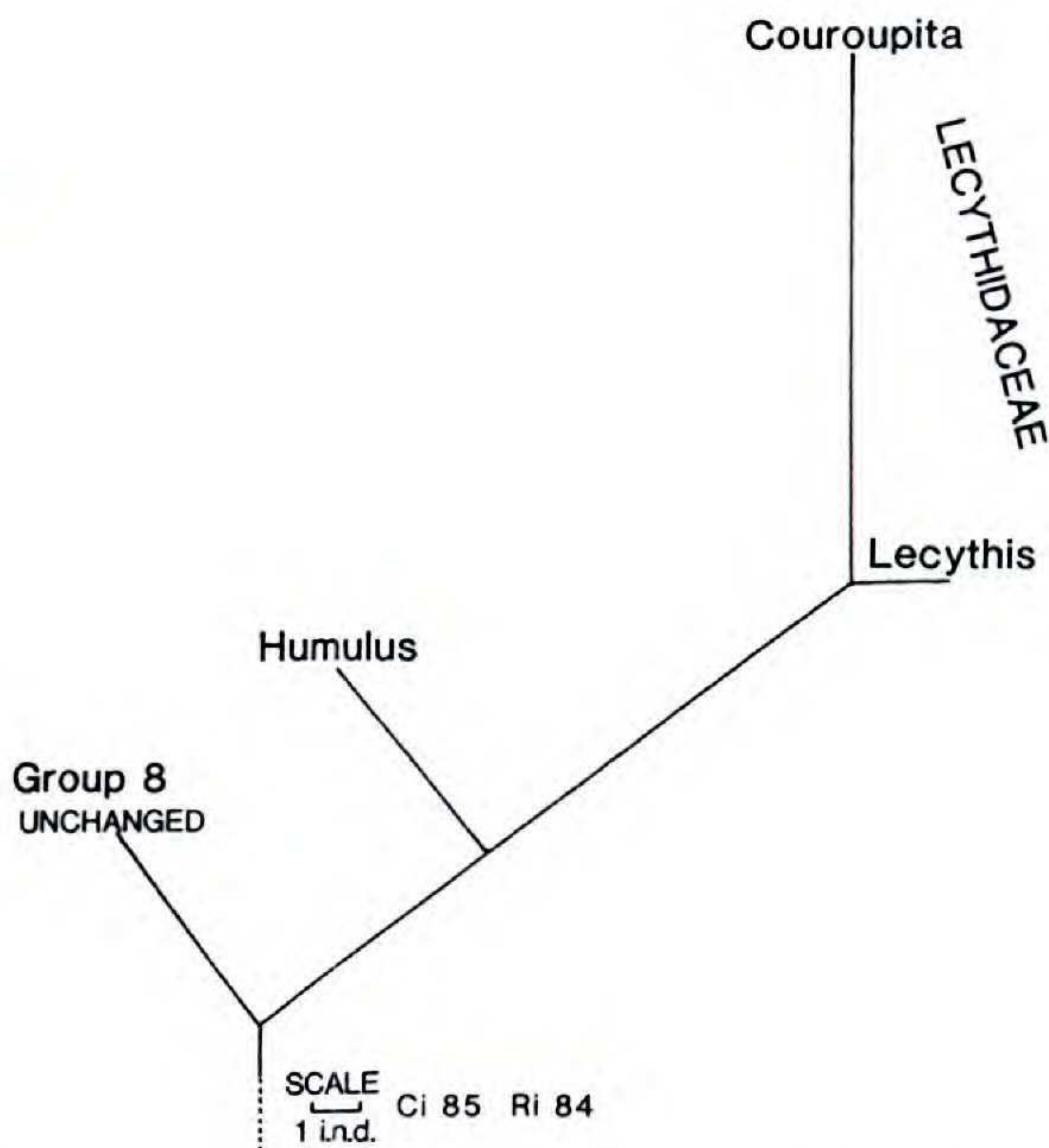


FIGURE 30. Group 8A. See Figure 11 for Group 8.

Group 12A. At all stages, Convolvulaceae and Polemoniaceae grouped separately from Solanaceae, the other member of Group 22. At the third stage, along with Polygonaceae, they clustered with Group 12 (Fig. 31). Polemoniaceae and Ericaceae are confused, but the other three families grouped appropriately.

Group 14A. The second stage tests suggested that the legumes should be divided between Caesalpinaceae on the one hand and Mimosaceae and Papilionaceae on the other; Caesalpinaceae clustered with Group 13. A series of Lake tests (see “First stage” above and Martin & Dowd, 1990) was therefore performed. These tests strongly indicated, first, that Caesalpinaceae was closer to Rosaceae than to either of the other two legume families and, second, that Mimosaceae and Papilionaceae were closer to other Groups (e.g., Connaraceae in Group 17, Chrysobalanaceae in Group 18A below) than were Caesalpinaceae and Rosaceae.

Other second stage tests had indicated that Proteaceae, *Coriaria*, *Crossosoma*, and Hamamelidaceae were also linked to the complex of Groups 13 and 14. The tree that resulted when these were

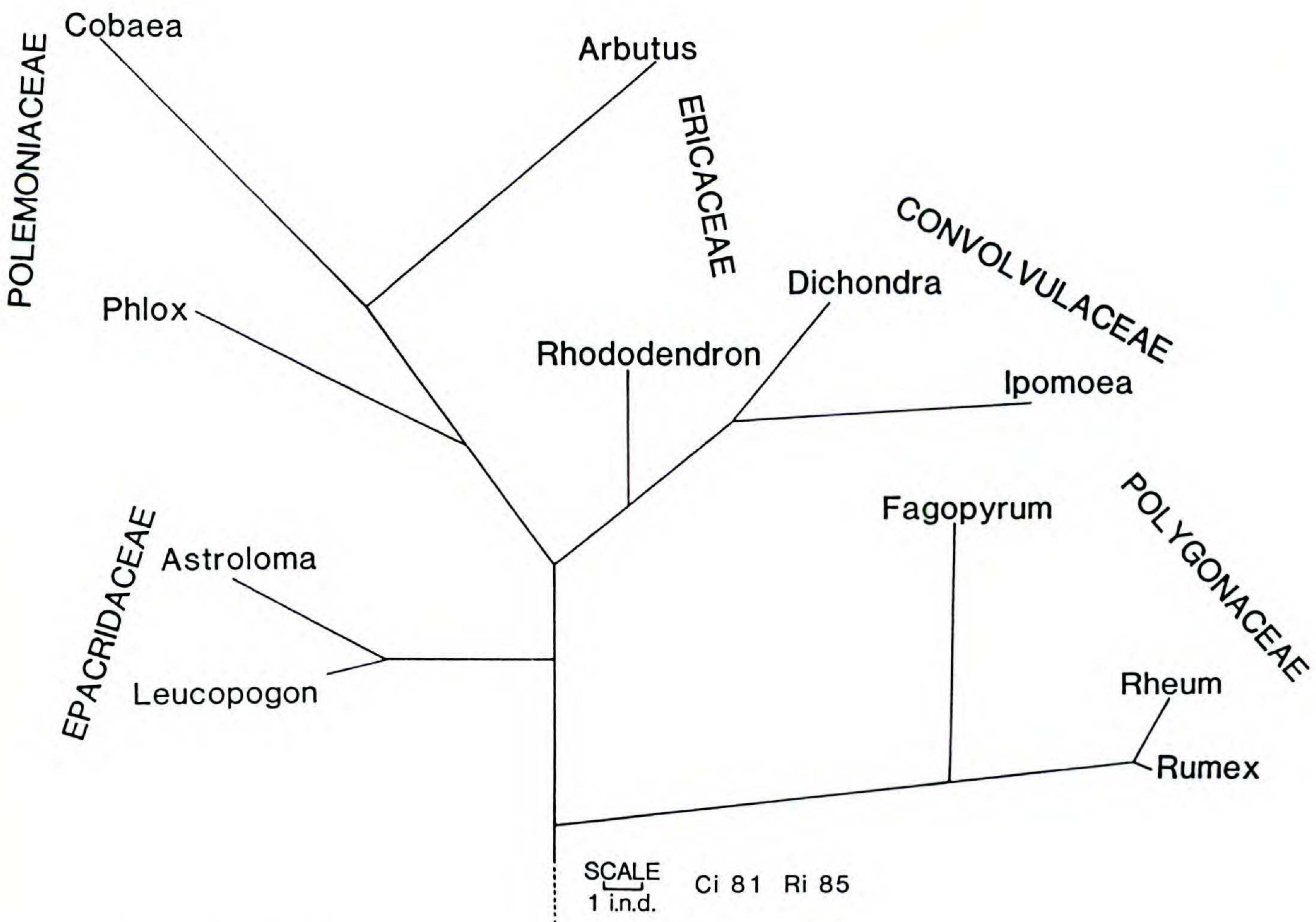


FIGURE 31. Group 12A.

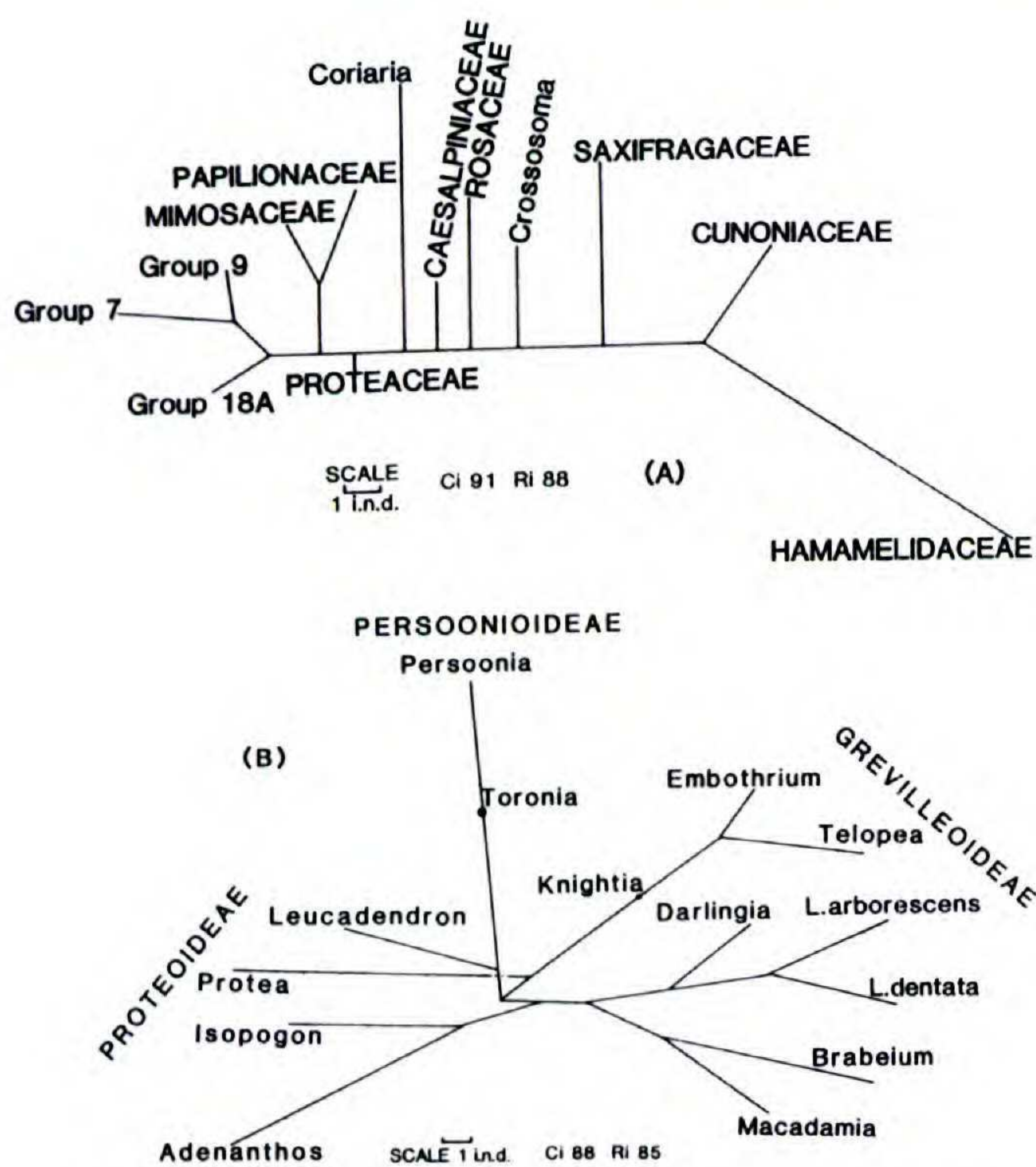


FIGURE 32. (A) Group 14A. For three legume families see Figure 17, and for Proteaceae see (B).

all analyzed together is shown in Figure 32a. The Proteaceae node was derived from Figure 32b, while the Mimosaceae–Papilionaceae node is node 3 of Figure 17a.

Group 18A. Third stage tests suggested that the Chrysobalanaceae and Vitaceae might cluster

with Group 18 and also with Group 25 (Apiaceae and Araliaceae). Incorporation of these (Fig. 33) does nothing to repair the previous (Fig. 21) disjunction of Haloragaceae while Rhizophoraceae s.l. remain apart from Anisophyllea.

Group 22A. With the other two families joining Group 12A (above), the Solanaceae were left as the sole representative.

Group 26. This new Group (Fig. 34a) consists of three families (Campanulaceae, Caprifoliaceae and Goodeniaceae), each with well-paired representatives. With them is Asteraceae, the node for which is derived from Figure 34b.

Group 27. This comprises the families Elaeagnaceae and Rhamnaceae, the members of which form pairs (Fig. 35).

Group 28. As noted below, *Buxus* does not pair with *Simmondsia*, which is sometimes placed in Buxaceae. While the latter clusters with Euphorbiaceae (Fig. 36), *Buxus* does not.

Group 29. The species of Hydrophyllaceae, Thymelaeaceae, and Valerianaceae form pairs in this new Group (Fig. 37).

TAXA THAT REMAIN UNPLACED

There are three families for which we have no acceptable hypothesis. (a) *Loasaceae*. It was un-

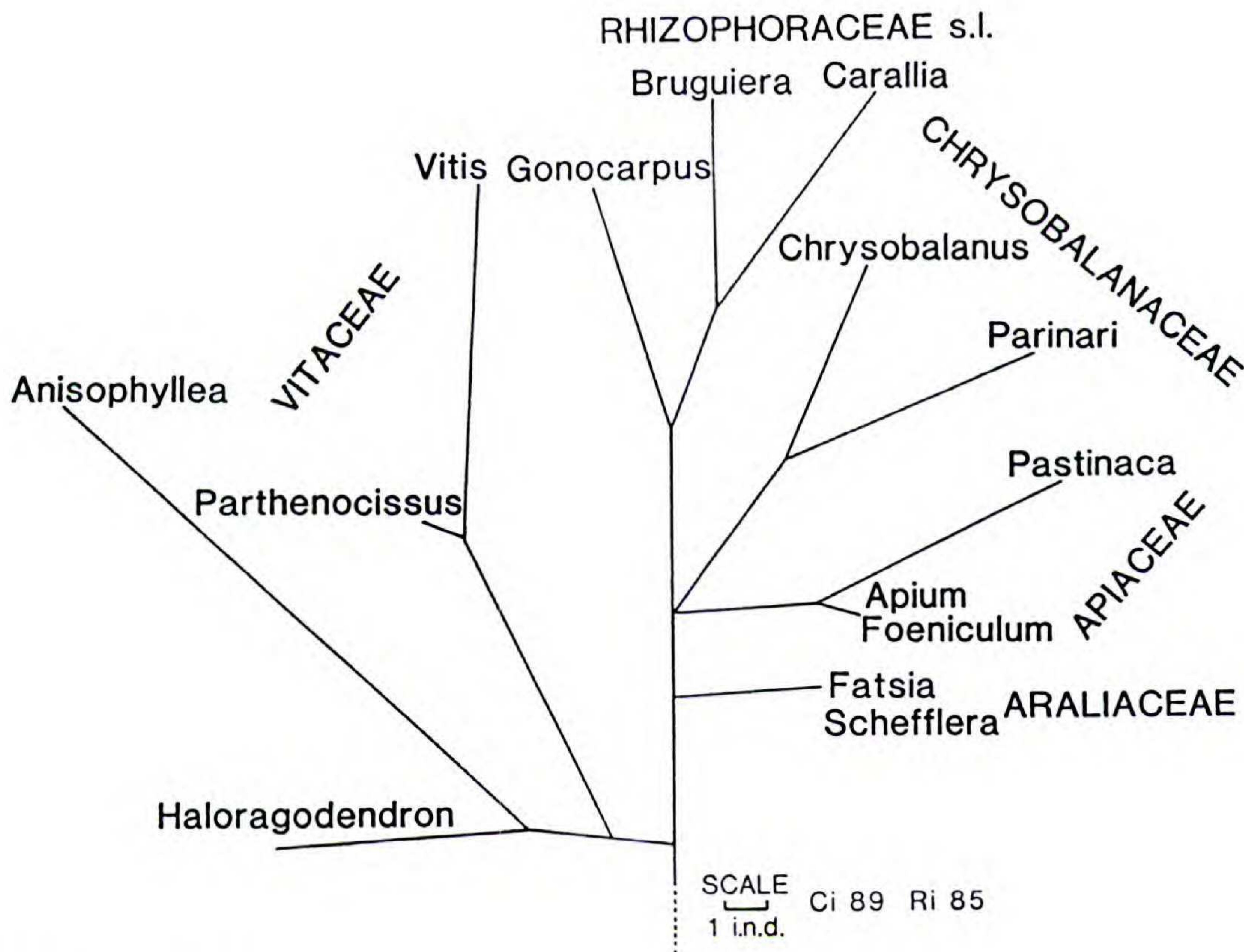


FIGURE 33. Group 18A.

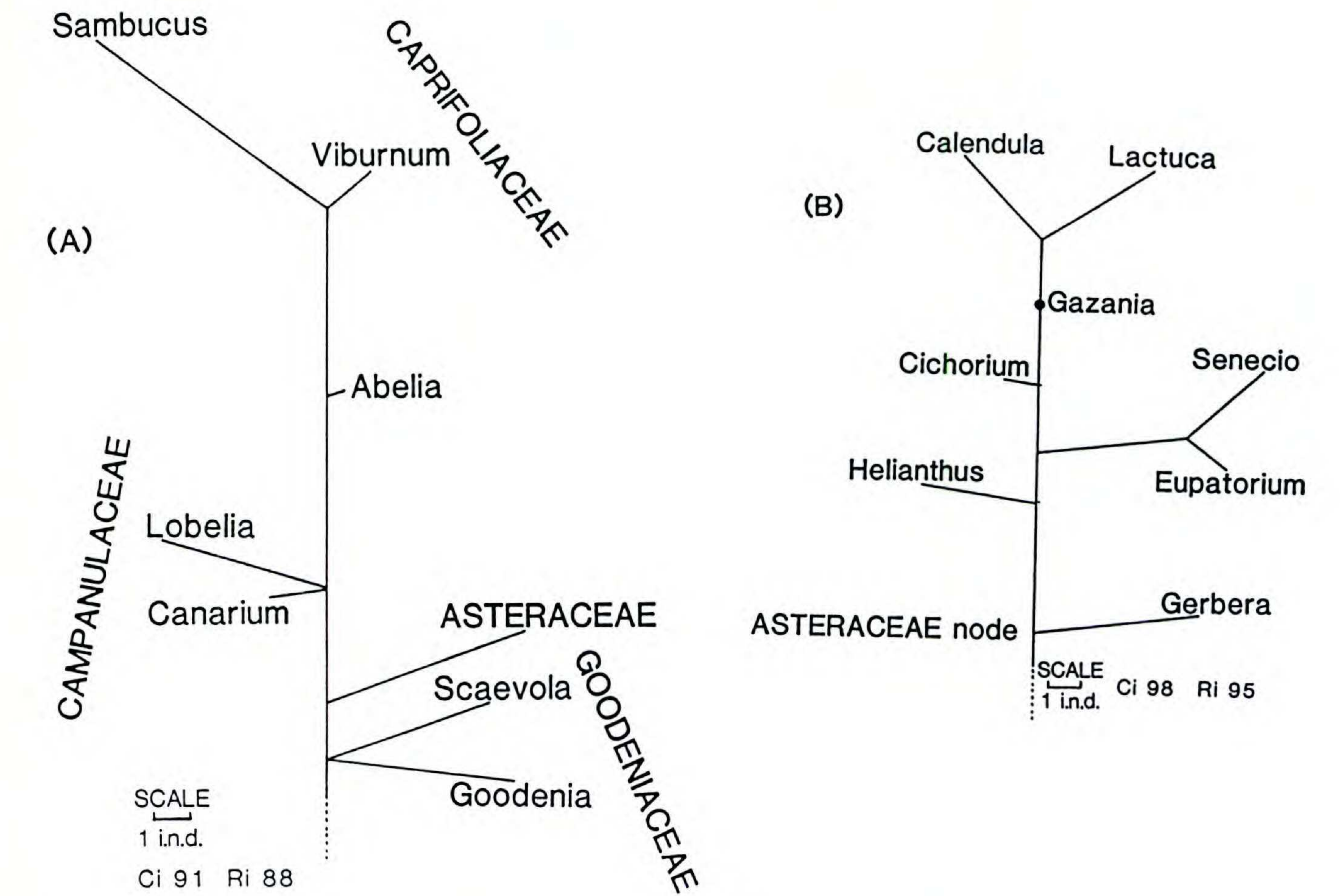


FIGURE 34. (A) Group 26 with node for Asteraceae from (B).

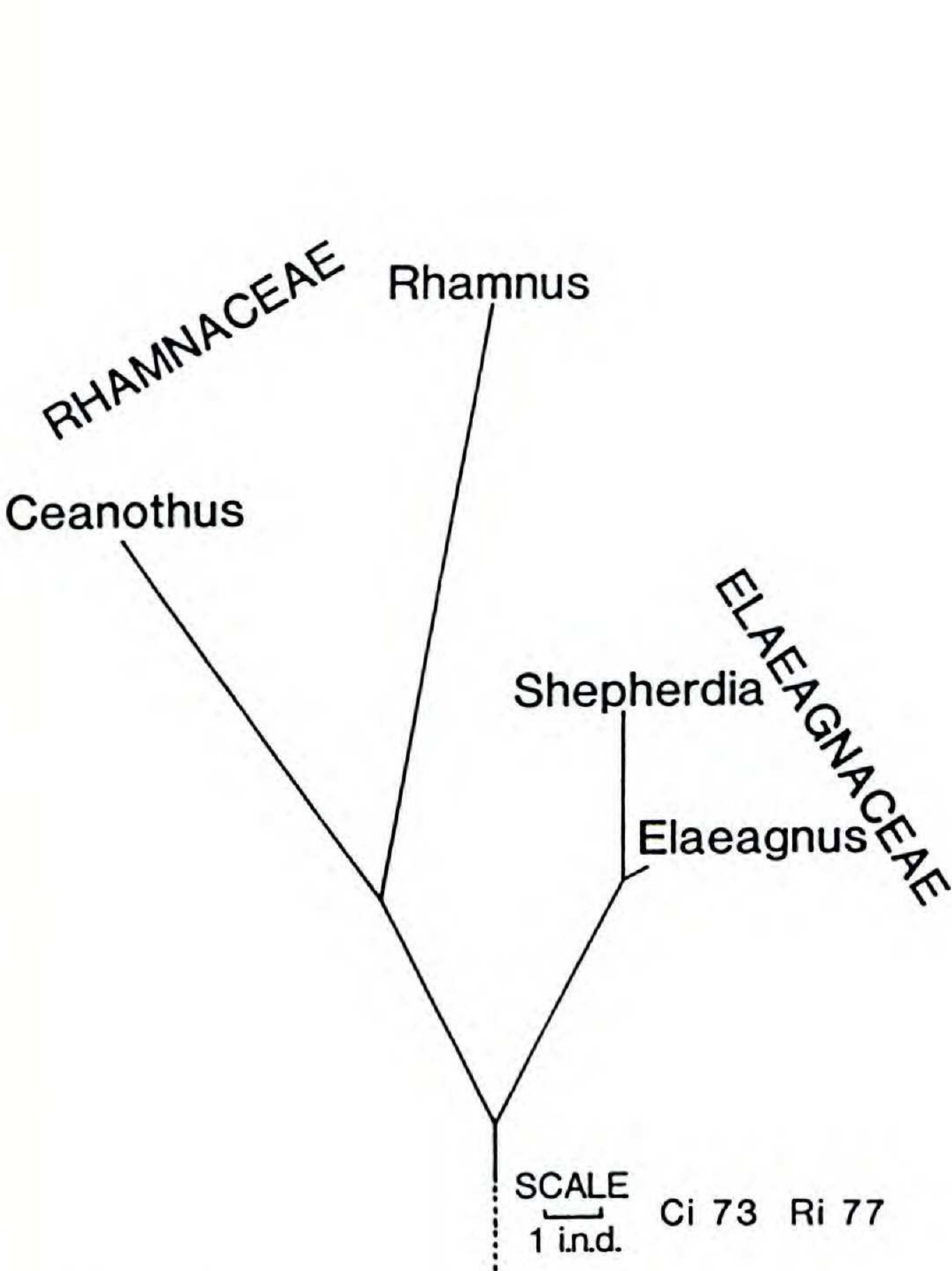


FIGURE 35. Group 27.

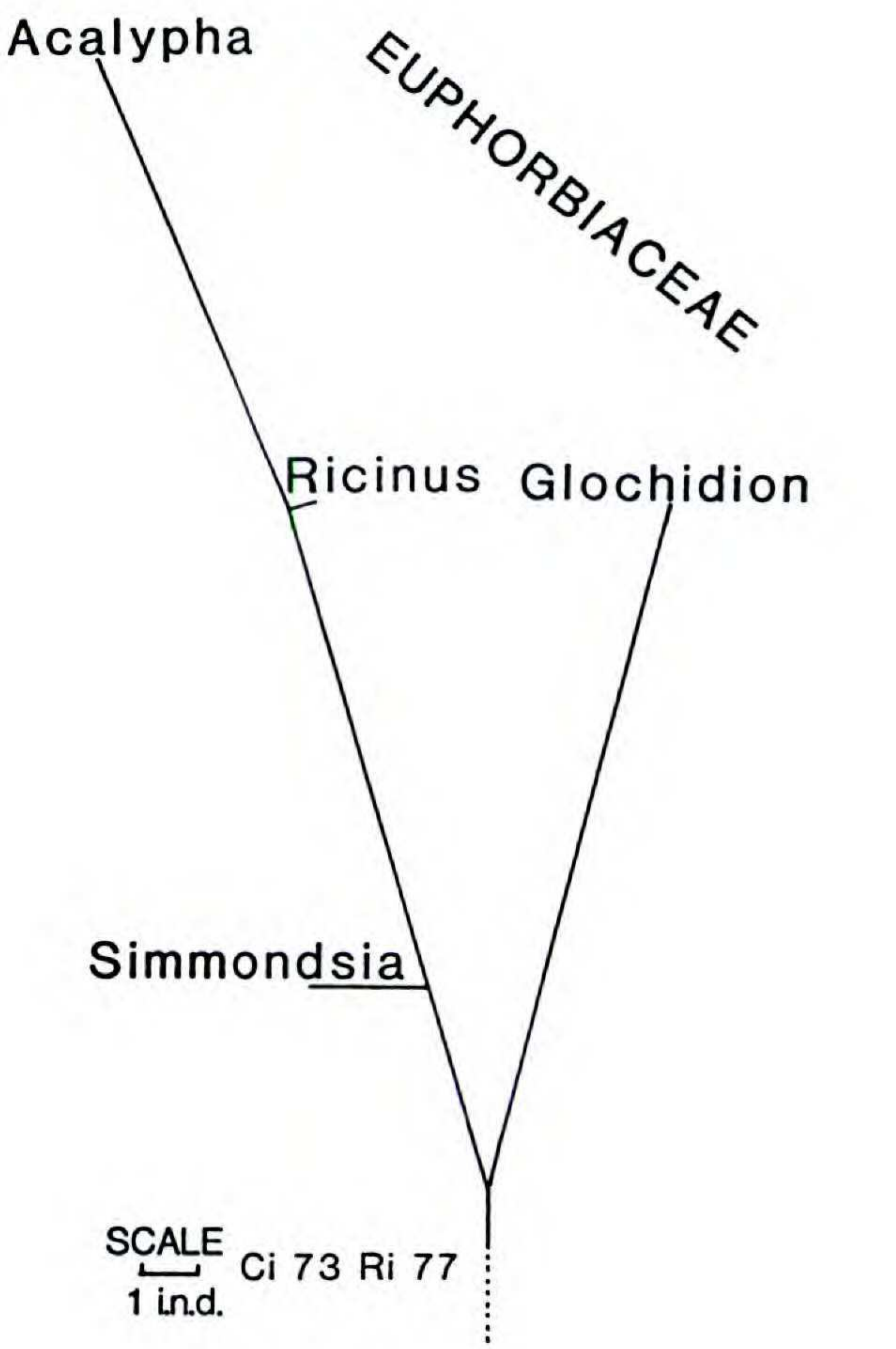


FIGURE 36. Group 28.

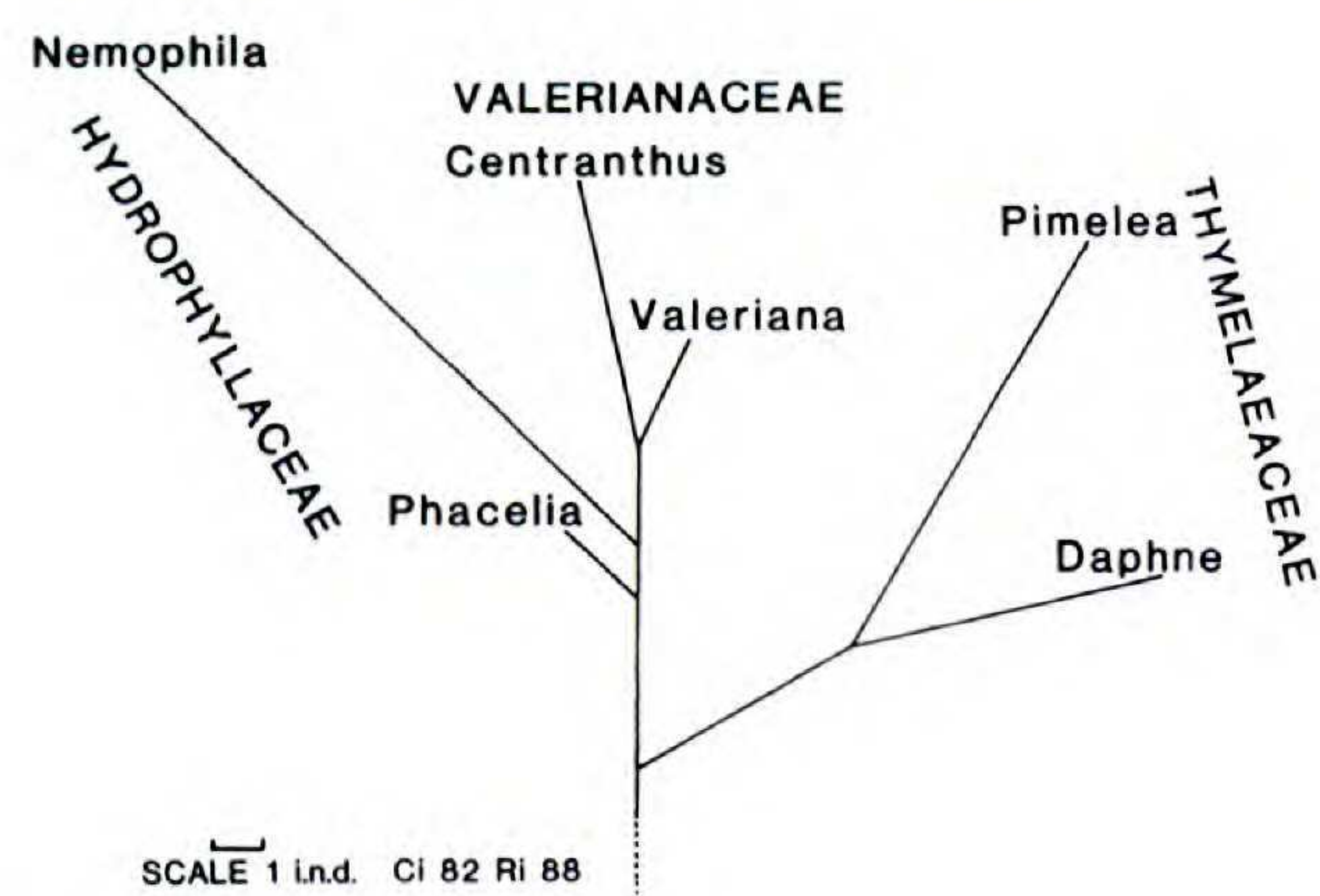


FIGURE 37. Group 29.

fortunate that we failed to obtain a sequence for *Eucnides bartonioides* because this left *Mentzelia* as a singleton and therefore with a “long edge” that joined unreliably. (b) *Plumbaginaceae*. Although the two representatives, *Limonium* and *Plumbago*, paired well, there remained a very long internode joining the family to the tree, and so we have left it unplaced. (c) *Buxaceae*. Originally both *Buxus* and *Simmondsia* were chosen as representatives of Buxaceae (s.l.), but they proved quite different and, since there was taxonomic opinion to support this, they were treated as such. Whereas

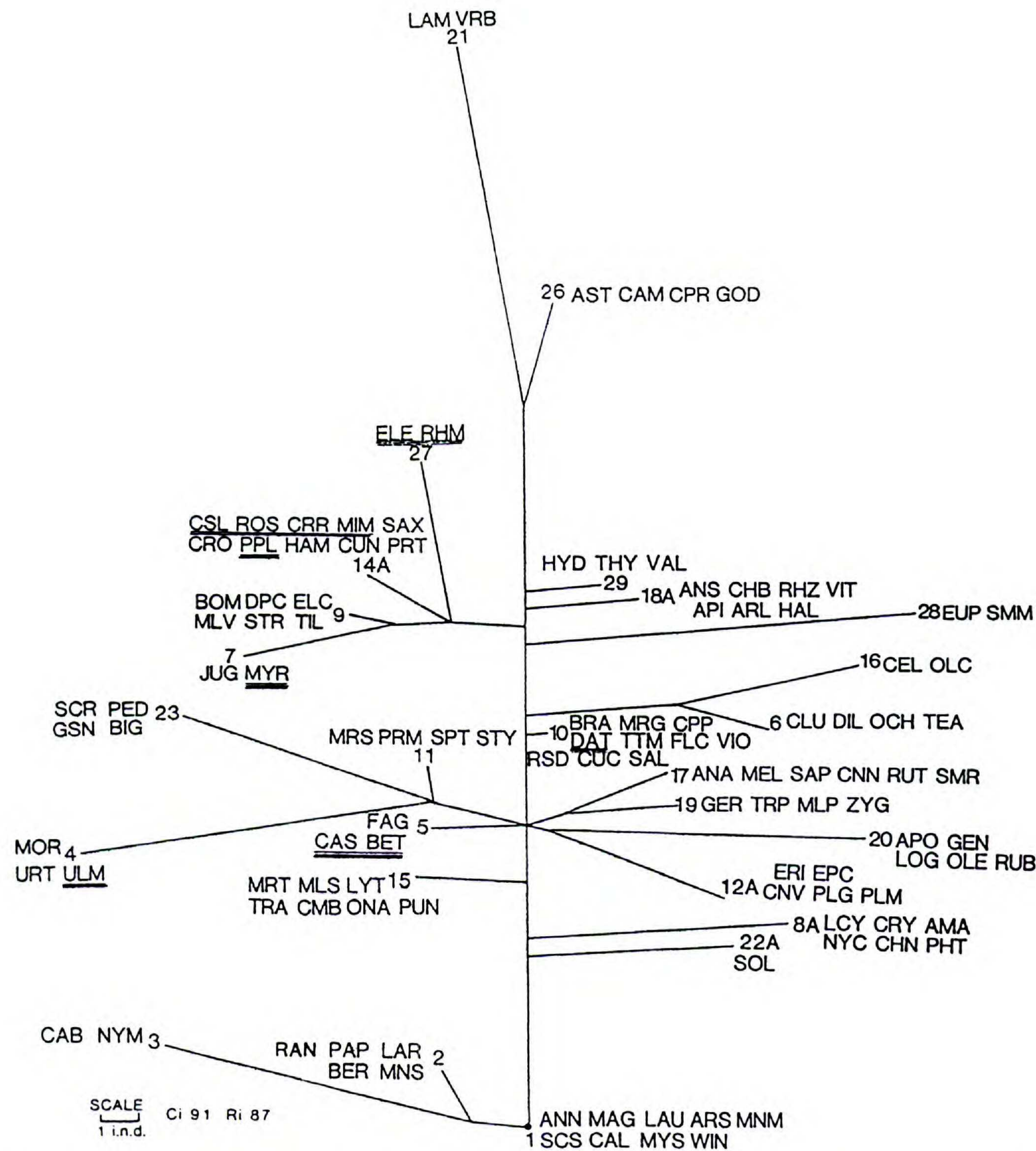


FIGURE 38. The overall tree for the dicotyledons. Groups are numbered and their constituent families indicated using the three-letter acronyms of Weber (1982), given in Table 2. Families in which nitrogen-fixation is known are underlined.

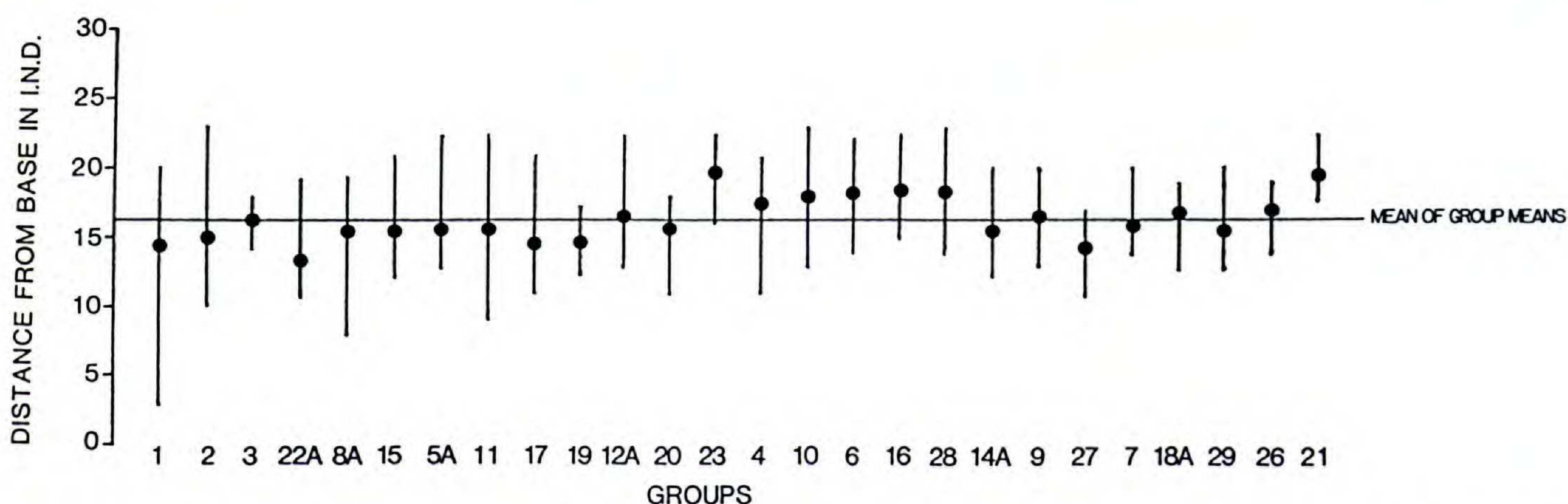


FIGURE 39. Groups are arranged along the X axis in the order that they depart from the trunk of the overall tree (Fig. 38). Solid dots are the mean distances of species of that Group from the angiosperm origin, and bars indicate the range from smallest to greatest.

Simmondsia grouped reasonably well with Euphorbiaceae, *Buxus* did not and remains unplaced.

FIFTH STAGE; THE SIMULTANEOUS ANALYSIS OF REVISED GROUPS

Initially, the nodes of all 26 revised Groups were analyzed using the option *mhennig* followed by *bb*, and the resulting tree was divided into a top, middle, and bottom section. Thus, with overlaps, each contained 14 taxa, a number that could be analyzed using the *ie* option. Fortunately, there was no confusion at the overlaps, and the three parts were fitted together to give the overall tree (Fig. 38).

DISCUSSION

THE RATE OF EVOLUTION AND THE AGE OF THE ANGIOSPERMS

In Figure 38, which shows the overall tree for the dicotyledons, there is a "trunk" from which branches depart at irregular intervals of up to 5 i.n.d. In Figure 39, we arranged Groups in the order that they branch from the trunk. For every species we measured the number of differences (in i.n.d.) between it and the base of the angiosperm tree (Fig. 1). For each Group we show the mean of these distances and also the range from smallest to greatest. The mean of all Groups is 16.2 i.n.d. We have also analyzed variance and shown that there is significant ($P < 0.001$) variation between Groups. Thus, although the difference between a slowly evolving Group such as Group 3 (mean 14.1) and a rapidly evolving Group such as Group 21 (19.7) is not great, it is probably real.

The age of the dicotyledons can be derived from the product of the mean number of differences of species from base and the rate of evolution. Since

Figure 6 suggests that the monocotyledons are derived from the dicotyledons this is also the age of the angiosperm. Martin and Dowd (1988) estimated the rate to be 1 i.n.d. in 14 Ma for a single evolutionary line. However, this estimate was based on members of the Fagaceae, Proteaceae, Solanaceae, and Winteraceae, all of which belong to Groups that evolve more slowly than average; their mean number of differences from base is 14.7 i.n.d. Thus, the inferred age of the angiosperms is $14 \times 14.7 = 205$ Ma, that is, at the beginning of the Jurassic. Crane et al. (1989) and Wolfe et al. (1989) have estimated the age of the angiosperms as 200 Ma. If the monocotyledons are indeed derived from the dicotyledons, there is good agreement.

THE RELIABILITY OF OUR TREES

The current limitations of computers and computing programs make it impossible to conduct a large phylogenetic analysis in a completely objective manner. Our first important deviation from objectivity has been accepting taxonomic opinion that species belong to the same family. Our second has been seeking a consensus in placing these into Groups.

The assumption of correct assignment to families is strongly supported by the correct pairing (or formation of clusters of three when appropriate) shown in the objectively derived figures of the final 26 Groups. Of the 95 families with two or three representatives, only 11 had disjunct representatives and, of these, at least four were families sensu lato with taxonomic opinions that they should really be split. These are the separation of *Humulus* from *Morus* in Group 4, of *Flindersia* from other Rutaceae (Group 17), of *Buxus* from *Simmondsia*

(Group 28), and of *Anisophyllea* from other Rhizophoraceae (Group 18). When it is further considered that one aberrant species can disrupt two families, we submit that the high proportion of correct grouping is strong evidence not only for the correctness of other taxonomy at this level but also of the soundness of our approach.

If our methods, while not perfect, are good at the level of placing taxa into families, is there any reason why they should not be equally acceptable at higher levels? We have investigated this with the assumption that the probability of errors will increase as internode lengths decrease. From each of the Group trees we have determined that the average length of internodes within families (restricting the measurements to families with only two correctly paired representatives) is 5.6 i.n.d., while the average length of internodes between families is 4.7. From the final tree showing the relationships of Groups (Fig. 38), the average length of internodes is 3.0. Thus, if our assumption is valid, the ratio 5.6:4.7:3.0 should reflect the reliability of arranging species within families, families within Groups, and Groups in the final tree. We suggest caution about accepting relationships as the taxonomic level increases.

There is no obvious reason why the ratio just reported should not be similar for other macromolecular sequences. However, with nucleic acid sequencing (see review by Palmer, 1988) the amount of information available might increase by an order of magnitude over that presented here; thus, even if internode lengths at the highest levels are still proportionately small, the probability of errors due to chance when using small numbers should diminish and lead to more decisive phylogenies.

THE VALUE OF THIS STUDY

We believe that the demarcation of plant taxa at all levels should be the prerogative of botanists with a broad background in taxonomy and that the same specialists are best suited to compare the results of this study, expressed as phylogenetic trees, with published phylogenies. Because we do not have that background, we resist the temptation to point out the similarities and differences that we perceive and to assess when our trees are likely to be incorrect. Our perceptions are likely to be unbalanced.

One difference between this phylogenetic study and most others is that it is repeatable. Without detracting from the value of published angiosperm phylogenies, they do seem to depend on the accumulated wisdom and experience of rare individuals whose relevant brain functions are not easily transmitted in entirety. On the other hand, anyone who follows our procedures should arrive at the same phylogenetic trees. More to the point, with improved analytical procedures it is possible that more acceptable endpoints may be reached.

We have avoided the word "conclusions" because we do not claim that this work is definitive. Rather it has led to new working hypotheses which, we hope, others will test with more extensive sampling and more data including much longer sequences. To such investigators our analytical method, whether perceived as successful or not, may be a useful example.

NATURAL SELECTION AND THE EVOLUTION OF RUBISCO-SSU

Under "General remarks about the sequences," we discussed heterogeneity within species and quoted the evidence of Pichersky et al. (1986) that natural selection acts to keep the amino acid sequence constant. Below we present other evidence for the importance of natural selection.

Under "Methods of Data Analysis," we discussed Lake's test, which is based only on transversions (mutations from a purine to a pyrimidine or vice versa) and ignores transitions (purine to purine or pyrimidine to pyrimidine). Lake (1987) quoted evidence (Brown et al., 1982) that in animal mitochondrial DNAs, transitions occur an order of magnitude more frequently than transversions. Zimmer et al. (1989) have found for higher plant cytoplasmic rRNA that, on average, transitions were twice as frequent as transversions with the lowest ratio in the most invariant regions. We have investigated this in 44 families of Groups 1 to 10 and have scored those amino acid changes within families that can be ascribed unequivocally to transversions and transitions. There were 123 transitions and 306 transversions, a proportion of 0.287 transitions and therefore quite different from the evidence just quoted.

We have considered each of the 61 codons in the genetic code and, assuming that each nucleotide can change to another with the same prob-

ability, calculated the frequencies of the four possibilities, i.e., transition causing amino acid change, transversion causing amino acid change, no amino acid change, and lethality (stop). Thus, for the two codons that determine phenylalanine the ratio of transitions to transversions is 0.25, for eight of the amino acids it is 0.33, and the ratio varies from 0.14 to 0.34 with an average of 0.2845. Averaging the 31 variable amino acids in the top line of Table 3 gives the ratio 0.268, which may be compared with the observed figure of 0.287. This suggests that, at the nonsilent positions, which are the only ones we are able to consider, there is close to randomness with respect to the occurrence of transitions and transversions.

We suggest that the large discrepancy between our result and the expectations from chemistry and nucleic acid sequencing is due partly to our inability to score silent substitutions and partly to the overwhelming importance of natural selection in determining the amino acid sequence of an important enzyme. Even though most nucleotide substitutions are presumably transitions, this has little effect on the final outcome, the amino acid sequence, on which natural selection can act.

Other evidence of strong natural selection comes from consideration of variation at positions like 8 and 9. At position 8, 84% of species have glycine and 15% asparagine. This substitution requires at least two nucleotide changes so, in the absence of selection, the single-change intermediates serine or lysine would often be expected, though they have not been observed. Similarly, at position 9, 56% of species have leucine and 28% lysine. Again, this is a two-nucleotide change, but the only single-change intermediates found are methionine and isoleucine, and these are much too rare to occur randomly. Apparently, glycine and asparagine are "adaptive peaks" at position 8 and leucine and lysine are at position 9. When positions 8 and 9 are considered together, there is a small excess over chance expectations of the combinations glycine-leucine and asparagine-lysine; these may be adaptive peaks because both combinations are found within Tiliaceae (Group 9), Papilionaceae (Group 14A), Apocynaceae (Group 20), Proteaceae (Group 14A), and different families of Group 15. Clearly, convergent evolution has occurred.

This last evidence suggests that adjacent positions influence one another, which is known. Another example is probably found in the Onagraceae,

the only family with N-terminal phenylalanine and, alongside it, asparagine, again only found in Onagraceae. *Solanum* species with the same rare substitutions at positions 15 and 21 are examples that the effect can extend further. Another example concerns positions 30 and 39, both of which are almost always either valine (V) or isoleucine (I). The frequencies within species of the four possible combinations (VV, VI, IV, II) indicate that the two positions evolve independently; nevertheless, they are different in the monocotyledons, with 67.5% isoleucine, and the dicotyledons, with 35.4% isoleucine. It is conceivable that monocotyledons are richer in isoleucine because they have a more efficient synthetic pathway for isoleucine so that, in the absence of other strong selective forces, the substitution of isoleucine for valine may be favored.

Despite the evidence that natural selection is acting strongly, there are few decisive changes, such as the change from proline to isoleucine at position 6 during the evolution of the monocotyledons. At positions 7 and 8, the combination tyrosine-asparagine occurs in the gymnosperms, Groups 1, 2, and 3, but in no other Groups, suggesting that these amino acids are primitive. However, the distinction between primitive and advanced is usually equivocal; for the following example, normal taxonomic criteria have been used to distinguish 58 primitive genera (those in the gymnosperms, Piperaceae, Nelumbonaceae, and Groups 1, 2, 3, and 5) from 67 advanced genera (those in Asteraceae, Campanulaceae, Goodeniaceae, Hydrophyllaceae, and Groups 10, 20, 21, 22, 23, and 24). At position 12, tyrosine occurred in 5% of primitive and 31% of advanced genera while at position 20 aspartic acid occurred in 5% of primitive and 43% of advanced genera. While admitting that the sampling is not entirely satisfactory, it appears that tyrosine at position 12 and aspartic acid at position 20 are advanced. However, the important point is that the divergence is so indecisive, the primitive amino acids phenylalanine at position 12 and proline at position 20 still occurring in the majority of genera in all advanced Groups.

If it is correct that natural selection acts strongly to determine the amino acid sequence of a protein, this could be important in considering "molecular evolutionary clocks." If the clock that is considered is derived from nucleic acid sequences, the rare event that is the basis of regression of number of

differences on time is nucleotide substitution, the most common form of mutation and not always subject to natural selection. If, however, the clock is derived from amino acid sequences, the rare

event is not mutation alone but its incorporation into the gene pool by natural selection. In both cases clocklike behavior is observed, but the rare event is different.

LITERATURE CITED

- BERRY-LOWE, S., T. D. MCKNIGHT, D. M. SHAH & R. B. MEAGHER. 1982. The nucleotide sequence, expression, and evolution of one member of a multigene family encoding the small subunit of ribulose-1,5-bisphosphate carboxylase in soybean. *J. Molec. Appl. Genet.* 1: 483-498.
- BROWN, W. M., E. M. PRAGER, A. WANG & A. C. WILSON. 1982. Mitochondrial sequences of primates: tempo and mode of evolution. *J. Molec. Evol.* 18: 225-239.
- CARPENTER, J. M. 1988. Choosing among multiple equally parsimonious cladograms. *Cladistics* 4: 291-296.
- CHAPMAN, M. S., S. W. SUH, P. M. G. CURMI, D. CASCIO, W. W. SMITH & D. S. EISENBERG. 1988. Tertiary structure of plant RuBisCO: domains and their contacts. *Science* 241: 71-74.
- CRANE, P. R., M. J. DONOGHUE, J. A. DOYLE & E. M. FRIIS. 1989. Angiosperm origins. *Nature (London)* 342: 131.
- CRONQUIST, A. 1976. The taxonomic significance of the structure of plant proteins; a classical taxonomist's view. *Brittonia* 28: 1-27.
- . 1981. *An Integrated System of Classification of Flowering Plants*. Columbia Univ. Press, New York.
- DAHLGREN, R. 1983. General aspects of angiosperm evolution and macrosystematics. *Nord. J. Bot.* 3: 119-149.
- FARRIS, J. S. 1988. HENNIG86 Reference: Version 1.5. James S. Farris, Port Jefferson Station, New York.
- GRUND, C., J. GILROY, T. GLEAVES, U. JENSEN & D. BOULTER. 1981. Systematic relationships of the Ranunculaceae based on amino acid sequence data. *Phytochemistry* 20: 1559-1565.
- HASLETT, B. G., A. YARWOOD, I. M. EVANS & D. BOULTER. 1976. Studies on the small subunit of fraction 1 protein from *Pisum sativum* L. and *Vicia faba* L. *Biochim. Biophys. Acta* 420: 122-132.
- HENDY, M. D. & D. PENNY. 1982. Branch and bound algorithms to determine minimal evolutionary trees. *Math. Biosci.* 59: 277-290.
- HEYWOOD, V. H. 1978. *Flowering Plants of the World*. Oxford Univ. Press, London.
- JOHNSON, D. 1987. More approaches to the travelling salesman guide. *Nature (London)* 330: 525.
- KNIGHT, S., I. ANDERSON & C. I. BRANDEN. 1989. Re-examination of the three-dimensional structure of the small subunit of RuBisCO from higher plants. *Science* 244: 702-705.
- LAKE, J. A. 1987. A rate-independent technique for analysis of nucleic acid sequences: evolutionary parsimony. *Molec. Biol. Evol.* 4: 167-191.
- LAMB, C. & L. FITZMAURICE. 1986. Tailoring crop improvement. *Nature (London)* 324: 414.
- MARTIN, G. C., J. KEEN, B. V. FORD-LLOYD & H. J. NEWBURY. 1987. Isolation and partial amino acid sequencing of the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from three species of orchid. *Pl. Sci.* 50: 233-237.
- MARTIN, P. G. 1979. Amino acid sequence of the small subunit of ribulose-1,5-bisphosphate carboxylase from spinach. *Austral. J. Pl. Physiol.* 6: 401-408.
- & J. M. DOWD. 1984a. The study of plant phylogeny using amino acid sequences of ribulose-1,5-bisphosphate carboxylase. III Addition of Malvaceae and Ranunculaceae to the phylogenetic tree. *Austral. J. Bot.* 32: 283-290.
- & ———. 1984b. The study of plant phylogeny using amino acid sequences of ribulose-1,5-bisphosphate carboxylase. IV Proteaceae and Fagaceae and the rate of evolution of the small subunit. *Austral. J. Bot.* 32: 291-299.
- & ———. 1984c. The study of plant phylogeny using amino acid sequences of ribulose-1,5-bisphosphate carboxylase. V Magnoliaceae, Polygonaceae and the concept of primitiveness. *Austral. J. Bot.* 32: 301-309.
- & ———. 1986a. Phylogenetic studies using protein sequences within the order Myrtales. *Ann. Missouri Bot. Gard.* 73: 442-448.
- & ———. 1986b. A phylogenetic tree for some monocotyledons and gymnosperms derived from protein sequences. *Taxon* 35: 460-475.
- & ———. 1986c. Is the gene for haemoglobin widespread among dicotyledons? Pp. 81-82 in W. Wallace & S. E. Smith (editors), *Proceedings of the Eighth Australian Nitrogen Fixation Conference*. Australian Institute of Agricultural Science, Parkville, Victoria.
- & ———. 1988. A molecular evolutionary clock for angiosperms. *Taxon* 37: 364-377.
- & ———. 1989. Phylogeny among the flowering plants as derived from amino acid sequence data. Pp. 195-204 in B. Fernholm, K. Bremer & H. Jornvall (editors), *The Hierarchy of Life*. Elsevier, Amsterdam.
- & ———. 1990. A protein sequence study of the dicotyledons and its relevance to the evolution of the legumes and nitrogen fixation. *Austral. Syst. Bot.* 3: 91-100.
- & A. C. JENNINGS. 1983. The study of plant phylogeny using amino acid sequences of ribulose-1,5-bisphosphate carboxylase. I Biochemical methods and the patterns of variability. *Austral. J. Bot.* 31: 395-409.
- , D. BOULTER & D. PENNY. 1985. Angiosperm phylogeny studied using sequences of five macromolecules. *Taxon* 34: 393-400.
- , J. M. DOWD & S. J. L. STONE. 1983. The study of plant phylogeny using amino acid sequences of ribulose-1,5-bisphosphate carboxylase. II The analysis of small subunit data to form phylogenetic trees. *Austral. J. Bot.* 31: 411-419.
- , ———, C. MORRIS & D. E. SYMON. 1986. The study of plant phylogeny using amino acid se-

- quences of ribulose-1,5-bisphosphate carboxylase. VI *Solanum* species from different continents. Austral. J. Bot. 34: 187–195.
- NAKANO, T., T. HASE & H. MATSUBARA. 1981. The complete amino acid sequence of parsley (*Petroselinum sativum*) ferredoxin. J. Biochem. 90: 1725–1730.
- PALMER, J. D., R. K. JANSEN, H. J. MICHAELS, M. W. CHASE & J. R. MAHHART. 1988. Chloroplast DNA variation and plant phylogeny. Ann. Missouri Bot. Gard. 75: 1180–1206.
- PENNY, D., M. D. HENDY & I. M. HENDERSON. 1987. Reliability of evolutionary trees. Cold Springs Harbour Symposium on Quantitative Biology 52: 857–862.
- PICHERSKY, E., R. BERNATSKY, S. D. TANKSLEY & A. R. CASHMORE. 1986. Evidence for selection as a mechanism in the concerted evolution of *Lycopersicon esculentum* (tomato) genes encoding the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase. Proc. Natl. Acad. U.S.A. 83: 3880–3884.
- RAMSHAW, J. A. M. 1982. Structures of plant proteins. Pp. 229–290 in D. Boulter & B. Parthier (editors), Nucleic Acids and Proteins in Plants 1. Encyclopaedia of Plant Physiology, NS Volume 14A. Springer, Berlin.
- ROY, H., A. VALERI, D. H. POPE, L. RUCKERT & K. A. COSTA. 1978. Small subunit contacts in ribulose-1,5-bisphosphate carboxylase. Biochemistry 17: 665–668.
- SCOGIN, R. 1981. Amino acid sequence studies and plant phylogeny. Pp. 229–290 in D. A. Young & D. S. Seidler (editors), Phytochemistry and Angiosperm Phylogeny. Praeger, New York.
- STROBAEK, S., G. C. GIBBONS, B. HASLETT, D. BOULTER & S. G. WILDMAN. 1976. On the nature of the polymorphism of the small subunit of ribulose-1,5-diphosphate carboxylase in the amphidiploid *Nicotiana tabacum*. Carlsberg Res. Commun. No. 41: 335–343.
- TAKHTAJAN, A. 1983. The systematic arrangement of dicotyledonous families. Pp. 180–201 in C. R. Metcalfe & L. Chalk (editors), Anatomy of the Dicotyledons, Vol. II (2nd Edition). Clarendon Press, Oxford.
- THORNE, R. F. 1983. Proposed new realignments in the angiosperms. Nordic J. Bot. 3: 85–117.
- WEBER, W. A. 1982. Mnemonic three-letter acronyms for the families of vascular plants: a device for more effective herbarium curation. Taxon 31: 74–88.
- WOLFE, K. H., M. GOUY, Y-W. YANG, P. M. SHARP & W-H. LI. 1989. Date of the monocot–dicot divergence estimated from chloroplast DNA sequence data. Proc. Nat. Acad. U.S.A. 86: 6201–6205.
- ZIMMER, E. A., R. K. HAMBY, M. L. ARNOLD, D. A. LEBLANC & E. L. THERIOT. 1989. Ribosomal RNA phylogenies and flowering plant evolution. Pp. 205–214 in B. Fernholm, K. Bremer & H. Jornvall (editors), The Hierarchy of Life. Elsevier, Amsterdam.